A COURSE

OF

PURE MATHEMATICS

CAMBRIDGE UNIVERSITY PRESS WAREHOUSE,
C. F. CLAY, MANAGER.
London: FETTER LANE, E.C.
Edinburgh: 100, PRINCES STREET.

Leipzig: F. A. BROCKHAUS.
Berlin: A. ASHER AND CO.
New York: G. P. PUTNAM'S SONS.
Bombay and Calcutta: MACMILLAN AND Co., LTD.

A COURSE

OF

PURE MATHEMATICS

BY

G. H. HARDY, M.A.

FELLOW AND LECTURER OF TRINITY COLLEGE, CAMBRIDGE

CAMBRIDGE:
At the University Press
1908

Cambridge:

PRINTED BY JOHN CLAY, M.A.

AT THE UNIVERSITY PRESS.

PREFACE.

THIS book has been designed primarily for the use of first year students at the Universities whose abilities reach or approach something like what is usually described as 'scholarship standard'. I hope that it may be useful to other classes of readers, but it is this class whose wants I have considered first. It is in any case a book for mathematicians: I have nowhere made any attempt to meet the needs of students of engineering or indeed any class of students whose interests are not primarily mathematical.

A considerable space is occupied with the discussion and application of the fundamental ideas of the Infinitesimal Calculus, Differential and Integral. But the general range of the book is a good deal wider than is usual in English treatises on the Calculus. There is at present hardly room for a new *Calculus* of an orthodox pattern. It is indeed not many years since there was urgent need of such a book, but the want has been met by the excellent treatises of Professors Gibson, Lamb, and Osgood, to all of which, I need hardly say, I am greatly indebted. And so I have included in this volume a good deal of matter that would find a place in any *Traité d'Analyse*, though in English books it is usually separated from the Calculus and classed as 'Higher Algebra' or 'Trigonometry'.

In the first chapter I have discussed in some detail the various classes of numbers included in the arithmetical continuum. I have not attempted to include any account of any purely arithmetical theory of irrational number, since I believe all such theories to be entirely unsuitable for elementary teaching. My aim in this chapter is a more modest one: I take the 'linear continuum' for granted and assume the existence of a definite number corresponding to each of its points; and all that I attempt to do is to

analyse and distinguish the various classes of numbers whose existence these assumptions involve.

Chapters II and III probably do not present many points of novelty. The account given in Chapter II of the most important classes of functions of x is more systematic and illustrated with much greater detail than is usual in English books. I have included, mainly for the sake of completeness, a certain amount of the elements of coordinate geometry of two and three dimensions: but I have, here and throughout the book, kept geometry in a strictly subordinate position and used it merely for purposes of illustration. I have also avoided any wealth of detail in connexion with the purely formal consequences of De Moivre's Theorem, and have devoted the space thus saved to the inclusion of a good deal of matter concerning vector analysis, bilinear transformation, and so on, which seemed to me likely to be more interesting and more useful as a preparation for Chapter X.

I have endeavoured to make Chapter IV one of the principal features of the book. The notion of a limit is one that has always presented grave difficulties to mathematical students even of great ability. It has been my good fortune during the last eight or nine years to have a share in the teaching of a good many of the ablest candidates for the Mathematical Tripos; and it is very rarely indeed that I have encountered a pupil who could face the simplest problem involving the ideas of infinity, limit, or continuity, with a vestige of the confidence with which he would deal with questions of a different character and of far greater intrinsic difficulty. I have indeed in an examination asked a dozen candidates, including several future Senior Wranglers, to sum the series $1 + x + x^2 + \dots$, and not received a single answer that was not practically worthless—and this from men quite capable of solving difficult problems connected with the curvature and torsion of twisted curves.

I cannot believe that this is due solely to the nature of the subject. There are difficulties in these ideas, no doubt: but they are not so great as many other difficulties inherent in mathematics that every young mathematician completely overcomes. The fault is not that of the subject or of the student, but of the text-book and the teacher. It is not enough for the latter, if he wishes to drive sound ideas on these points well into the mind of his pupils,

to be careful and exact himself. He must be prepared not merely to tell the truth, but to tell it elaborately and ostentatiously. He must drill his pupils in 'infinity' and 'continuity', with an abundance of written exercises and examples, as he drills them at present in poles and polars or symmetric functions or the consequences of De Moivre's theorem. Then and only then he may hope that accurate thought in connexion with these matters will become an integral part of their ordinary mathematical habit of mind. It is this conviction that has led me to devote so much space to the most elementary ideas of all connected with limits, to be purposely diffuse about fundamental points, to illustrate them by so elaborate a system of examples, and to write a chapter of fifty pages without advancing beyond the ordinary geometrical series.

It is not necessary for me to say much about the general plan of the next four chapters. The two chapters on the Calculus are no doubt more difficult than the rest of the book. I have perhaps been inconsistent in the standards that I have adopted: but I have been influenced by the feeling that I shall have few readers who will not already have acquired some familiarity with the technique of the Calculus from other sources. I felt this particularly when I was writing the sections on integration. I also felt that the student is apt to carry away from the books in general use the quite mistaken impression that all methods of integration are essentially of a tentative and haphazard character. I have therefore deliberately given an account of the theory more systematic and general than would be suitable for a normal first course in the Calculus.

Chapters IX and X are devoted to the theory of the logarithm and exponential, starting from the definition of the logarithm as an integral. It was the desire to write an elementary account of this theory that originally led me to begin the book, and I have generally decided my choice of what was to be included in the earlier chapters by a consideration of what theorems would be wanted in the last two.

I regard the book as being really elementary. There are plenty of hard examples (mainly at the ends of the chapters): to these I have added, wherever space permitted, an outline of the solution. But I have done my best to avoid the inclusion of anything that involves really difficult ideas. For instance, I make no

use of the 'principle of convergence': uniform convergence, double series, infinite products, are never alluded to: and I prove no general theorems whatever concerning the inversion of limit-operations—I never even define $\frac{d^2f}{dx\,dy}$ and $\frac{d^2f}{dy\,dx}$. In the last two chapters I have occasion once or twice to integrate a power-series, but I have confined myself to the very simplest cases and given a special discussion in each instance. Anyone who has read this book will be in a position to read with profit Mr Bromwich's *Infinite Series*, where a full and adequate discussion of all these points will be found.

It will be found that certain classes of theorems and examples that are prominent in many English books are here conspicuous by their absence. I may refer particularly to the standard theorems concerning the expression of the trigonometrical functions as infinite products or series of partial fractions, and to that familiar type of example the gist of which lies in the 'picking out of coefficients' from some combination of infinite series. The proofs of these results depend upon general theorems that seemed to me intrinsically too difficult to be included in a book professing to be at the same time rigorous and elementary: and I am on the whole of opinion that, if any proposition is too difficult to be proved properly, its statement and application had better be postponed. I am well aware that there is much to be said on the opposite side. A very plausible case can be made out for the habitual exercise of the student in the application of results whose proof is too difficult for his full comprehension. But I have found that I cannot myself write a book on those lines: nor am I fully convinced that such exercise is either necessary or desirable. After all there are plenty of theorems which are *not* too difficult to prove: and, if anyone believes that a sufficient variety of analytical training cannot be based upon them, I hope that my collections of Miscellaneous Examples may do something to convince him. I may say that it is only in these collections that examples of the character of 'problems' will be found. The sets of examples inside each chapter consist either of perfectly straightforward applications of the preceding 'book-work', or of summaries of parts of the theory for which there was no room in the text. They include many important theorems, some indeed to which reference is frequently

made later in the book. No one can be more convinced than I am of the value of 'examples' designed merely to train the student's memory and powers of manipulation: but I see no reason why all examples should necessarily be trivial. I trust, however, that readers will not find it irritating to be referred back from the middle of a section in large type to an example in an earlier chapter. My decision as to whether a result should appear in the text or in the examples has always been based upon the relation that it bears to the general theorems in connexion with which it is first proved rather than upon the amount of use that is made of it later on.

I have throughout laid particular stress upon points that do not seem to me to be emphasized sufficiently in the text-books in general use, and passed rapidly over others that are of equal importance but stand in no such danger of neglect. Here again I have been influenced by the consideration that this book is likely to be used in conjunction with others rather than as a first text-book in any particular subject.

There are two respects in which I have diverged from the usually accepted notation and that seem of sufficient importance to be noticed here. I have entirely rejected the index notation for inverse functions ($\cos^{-1} x$, $\cosh^{-1} x$) in favour of the usual Continental notation ($\operatorname{arc} \cos x$, $\operatorname{arg} \cosh x$ or $\operatorname{arg} \operatorname{ch} x$). And I have followed Mr Leathem and Mr Bromwich in always writing

$$\lim_{n \to \infty}, \quad \lim_{x \to \infty}, \quad \lim_{x \to a}$$

and not $\lim\limits_{n=\infty}$, $\lim\limits_{x=\infty}$, $\lim\limits_{x=a}$. This last change seems to me one of considerable importance, especially when '∞' is the 'limiting value'. I believe that to write '$n = \infty$, $x = \infty$' (as if anything ever were 'equal to infinity'), however convenient it may be at a later stage, is in the early stages of mathematical training to go out of one's way to encourage incoherence and confusion of thought concerning the fundamental ideas of analysis.

The word 'quantity' occurs occasionally in the earlier chapters. It should be in each case altered to 'number'. Unfortunately I arrived at the decision never to use the term 'quantity' only after the earlier sheets had been passed for press.

The books to which I am most indebted (besides the treatises on the Calculus already mentioned) are Mr Bromwich's *Infinite*

Series and M. J. Tannery's *Leçons d'Algèbre et d'Analyse*. I must also acknowledge my obligations to a number of friends who have been kind enough to assist me in the preparation of the book. Mr Bromwich has read the whole of it (except Chapter III) either in manuscript or in proof, and a good deal of it twice; and I am indebted to him for corrections and suggestions on almost every page. Mr Berry read Chapters I, II, III, IX and X in manuscript, Professor J. E. Wright Chapters I, II, and III, and Dr Whitehead Chapters I and IV, and all gave me much valuable advice. In particular the earlier part of Chapter IV has been practically rewritten in consequence of Dr Whitehead's suggestions. I have also changed a good deal of Chapter VI in consequence of suggestions received from Dr Askwith. My thanks are also due to Messrs H. W. Turnbull and E. H. Neville, of Trinity College, who have between them read all the proofs and verified the examples: to the latter I am additionally indebted for the figures that appear in the Miscellaneous Examples to Chapter X. Finally I must express my gratitude to the readers and officials of the University Press for their close attention and unfailing courtesy.

G. H. HARDY.

TRINITY COLLEGE, CAMBRIDGE,
September 1908.

CONTENTS.

CHAPTER I.

REAL VARIABLES.

CHAPTER II.

FUNCTIONS OF REAL VARIABLES.

CHAPTER III.

COMPLEX NUMBERS.

CHAPTER IV.

LIMITS OF FUNCTIONS OF A POSITIVE INTEGRAL VARIABLE.

CHAPTER V.

LIMITS OF FUNCTIONS OF A CONTINUOUS VARIABLE: CONTINUOUS AND DISCONTINUOUS FUNCTIONS.

CHAPTER VI.

DERIVATIVES AND INTEGRALS.

CHAPTER VII.

ADDITIONAL THEOREMS IN THE DIFFERENTIAL AND INTEGRAL CALCULUS.

CHAPTER VIII.

THE CONVERGENCE OF INFINITE SERIES AND INFINITE INTEGRALS.

CHAPTER IX.

THE LOGARITHMIC AND EXPONENTIAL FUNCTIONS OF A REAL VARIABLE.

CHAPTER X.

THE GENERAL THEORY OF THE LOGARITHMIC, EXPONENTIAL, AND CIRCULAR FUNCTIONS.

CHAPTER I.

REAL VARIABLES.

1. The aggregate of rational numbers and their representation on a straight line. On a straight line L, produced indefinitely in both directions, we take a segment A_0A_1 of any length. We call A_0 *the origin*, or *the point* 0, and A_1 *the point* 1.

We now mark off a series of points

$$\ldots\ A_{-m-1},\ A_{-m},\ \ldots,\ A_{-1},\ A_0,\ A_1,\ \ldots,\ A_n,\ \ldots$$

along L, so that

$$\ldots = A_{-m-1}A_{-m} = \ldots = A_{-1}A_0 = A_0A_1 = \ldots,$$

each segment being measured from left to right along L.

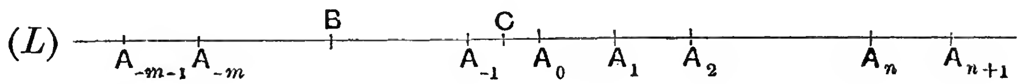


FIG. 1.

Then $$\frac{A_0A_n}{A_0A_1} = n \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1),$$

if n is any positive integer.

We will now agree that *length* is to be regarded as *a magnitude capable of sign*, positive if the length is measured in one direction along L (e.g. from B to C) and negative if measured in the other (from C to B), so that $CB = -BC$. We take the positive direction for the measurement of length to be from left to right.

Then $$\frac{A_0A_{-n}}{A_0A_1} = -\frac{A_{-n}A_0}{A_0A_1} = -n.$$

Hence the equation (1) is true for all integral values of n, positive or negative.

For the sake of uniformity we adopt the convention that (1) is also true when $n = 0$, in which case it reads

$$\frac{A_0A_0}{A_0A_1} = 0.$$

That is to say, we agree to regard BB, which is not, properly speaking, a segment at all, as a segment of no length.

Now let us take any positive proper fraction in its lowest terms, for example p/q, where p and q are positive integers without any common factor, and $p < q$. We divide A_0A_1 into q equal parts by points of division which it is natural to denote by

$$A_0,\ A_{1/q},\ A_{2/q},\ \ldots,\ A_{p/q},\ \ldots,\ A_{(q-1)/q},\ A_1.$$

It is evident that

$$\frac{A_0A_{p/q}}{A_0A_1} = \frac{p}{q} \quad \text{.........................(2).}$$

We thus obtain points on the line L corresponding to all such proper fractions p/q.

Any improper fraction may be expressed in the form $n + (p/q)$, where n is a positive integer and p/q a proper fraction. If we take a point $A_{n+(p/q)}$ such that $A_nA_{n+(p/q)} = A_0A_{p/q}$, it is evident that $\dfrac{A_0A_{n+(p/q)}}{A_0A_1} = n + \dfrac{p}{q}$: and if we thus find points $A_{n+(p/q)}$ corresponding to all possible positive values of n, p, and q, we shall have a point A_f corresponding to all possible positive integral or fractional values of f, and such that

$$\frac{A_0A_f}{A_0A_1} = f \quad \text{.........................(3).}$$

Finally, if $-f$ is a negative fraction, proper or improper, we take A_{-f} so that $A_{-f}A_0 = A_0A_f$, or

$$\frac{A_0A_{-f}}{A_0A_1} = -\frac{A_0A_f}{A_0A_1} = -f.$$

Thus we are able to determine a point A_r corresponding to *any integral or fractional value of* r, *positive or negative*, and such that

$$\frac{A_0A_r}{A_0A_1} = r \quad \text{.........................(4).}$$

If we take, as is natural, the length A_0A_1 as our unit of length, so that $A_0A_1 = 1$, the equation (4) becomes

$$A_0A_r = r \quad \text{.........................(5).}$$

DEFINITIONS. *Any fraction $r = p/q$, where p and q are positive or negative integers, is called a* **rational number.**

The points A_r of the line L, which correspond to the rational numbers r in the manner explained above, are called the **rational points** *of the line.*

We can suppose (i) that p and q have no common factor, as if they have a common factor we can divide each of them by it, and (ii) that q is positive, since

$$p/(-q) = (-p)/q, \quad (-p)/(-q) = p/q.$$

The notion of a rational number obviously includes as a particular case that of an integer, since any integer may be expressed as a fraction whose denominator is unity.

Examples I. 1. If r and s are rational numbers, $r+s$, $r-s$, rs, and r/s are rational numbers, unless in the last case $s=0$ (when r/s is of course meaningless).

2. If P and Q are rational points, and PQ is divided into any number of equal parts, each of the points of division is a rational point.

3. If λ, m, and n are positive rational numbers, $\lambda(m^2 \sim n^2)$, $2\lambda mn$, and $\lambda(m^2+n^2)$ are positive rational numbers. Hence show how to determine any number of right-angled triangles the lengths of all of whose sides are rational.

4. Any terminated decimal represents a rational number whose denominator contains no factors other than 2 or 5. Conversely, any such rational number can be expressed, and in one way only, as a terminated decimal.

[The general theory of decimals will be considered in Chap. IV.]

5. The positive rational numbers may be arranged in the form of a simple series as follows:

$$\tfrac{1}{1}, \tfrac{2}{1}, \tfrac{1}{2}, \tfrac{3}{1}, \tfrac{2}{2}, \tfrac{1}{3}, \tfrac{4}{1}, \tfrac{3}{2}, \tfrac{2}{3}, \tfrac{1}{4}, \ldots$$

Show that p/q is the $[\frac{1}{2}(p+q-1)(p+q-2)+q]$th term of the series.

[In this series every rational number is repeated indefinitely. Thus 1 occurs as $\frac{1}{1}, \frac{2}{2}, \frac{3}{3}, \ldots$. We can of course avoid this by omitting every number which has already occurred in a simpler form, but then the problem of determining the precise position of p/q becomes more complicated.]

2. Irrational numbers. If the reader will mark off on the line all the points corresponding to the rational numbers whose denominators are 1, 2, 3, ... in succession, he will readily convince himself that he can cover the line with rational points as closely as he likes. We can state this more precisely as follows: *if we take any segment BC on L, we can find as many rational points as we please on BC.*

Suppose, for example, that BC falls within the segment A_1A_2. It is evident that if we choose a positive integer k so that

$$k \, . \, BC > A_0A_1 \qquad \text{.........................(1)}$$

and divide A_1A_2 into k equal parts, at least one of the points of division (say P) must fall *inside* BC, without coinciding with either B or C. For if this were not so BC would be entirely included in one of the k parts into which A_1A_2 has been divided, which contradicts the supposition (1). Thus at least one rational point P lies between B and C. But then we can find another such point Q between B and P, another between B and Q, and so on indefinitely; i.e., as we asserted above, we can find as many as we please. We may express this by saying that BC includes *infinitely many* rational points.

From these considerations the reader might be tempted to infer that these rational points account for all the points of the line, i.e. that *every* point on the line is a rational point. And it is certainly the case that if we imagine the line as being made up solely of the rational points, all other points (if any such there be) being imagined to be eliminated, the figure which remained would possess most of the properties which common sense attributes to the straight line, and would, to put the matter roughly, look and behave very much like a line.

There is, however, good reason for supposing that *there are on the line points which are not rational points.*

Let us look at the matter for a moment with the eye of common sense, and consider some of the properties which we may reasonably expect a straight line to possess if it is to satisfy the idea which we have formed of it in elementary geometry.

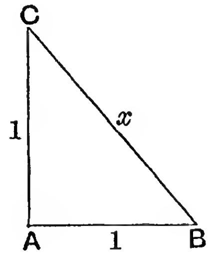
The straight line must be composed of points, and any segment of it by all the points which lie between its end points. With

any such segment must be associated a certain entity called its *length*, which must be a *quantity* capable of *numerical measurement* in terms of any standard or unit length, and these lengths must be capable of combination with one another according to the ordinary rules of algebra by means of addition or multiplication. Again, it must be possible to construct a line whose length is the sum or product of any two given lengths. If the length PQ, along a given line, is a, and the length QR, along the same straight line, is b, the length PR must be $a+b$. Moreover, if the lengths OP, OQ, along one straight line, are 1 and a, and the length OR along another straight line is b, and if we determine the length OS by Euclid's construction (Euc. VI. 12) for a fourth proportional to the lines OP, OQ, OR, this length must be ab, the algebraical fourth proportional to 1, a, b. And it is hardly necessary to remark that the sums and products thus defined must obey the ordinary laws of algebra, such as

$$a+b=b+a, \quad a+(b+c)=(a+b)+c, \quad ab=ba,$$

and so on. The lengths of our lines must also obey a number of obvious laws concerning *inequalities* as well as equalities. Thus if A, B, C are three points lying along L from left to right, we must have $AB < AC$, and so on. Finally it must be possible, on our fundamental line L, to find a point P such that A_0P is equal to any segment whatever taken along L or along any other straight line.

Now it is very easy, by means of various elementary geometrical constructions, to construct a length x such that $x^2=2$. For example, we may construct an isosceles right-angled triangle



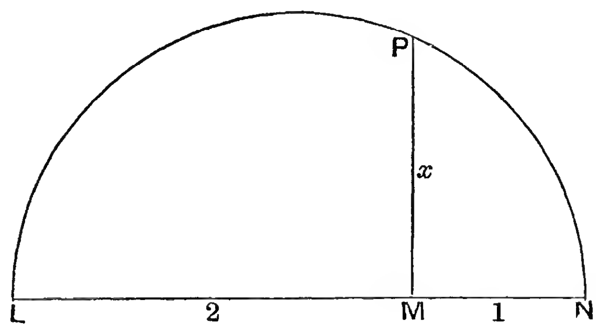


FIG. 2.

ABC such that $AB=AC=1$. Then if $BC=x$, $x^2=2$. Or we may determine the length x by means of Euclid's construction

(Euc. VI. 13) for a mean proportional to 1 and 2, as indicated in the figure.

It follows that there must be a point P on L such that

$$A_0P = x, \quad x^2 = 2.$$

But it is easy to see that *there is no rational number such that its square is* 2. In fact we may go further and say that there is no rational number whose square is m/n, where m/n is any positive fraction in its lowest terms, unless m and n are both perfect squares.

For suppose, if possible, that

$$\frac{p^2}{q^2} = \frac{m}{n},$$

p having no factor in common with q, and m no factor in common with n. Then

$$np^2 = mq^2.$$

Every factor of q^2 must divide np^2, and as p and q have no common factor, every factor of q^2 must divide n. Hence $n = \lambda q^2$, where λ is an integer. But this involves $m = \lambda p^2$: and as m and n have no common factor, λ must be unity. Thus $m = p^2$, $n = q^2$, as was to be proved.

We are thus led to believe in the existence of a point P, not one of the rational points already constructed, and such that $A_0P = x$, $x^2 = 2$; and (as the reader will remember from elementary algebra) we write $x = \sqrt{2}$. And if Q is the point such that $QA_0 = A_0P$, we write $A_0Q = -\sqrt{2}$.

The following alternative proof that $\sqrt{2}$ cannot be rational is interesting.

Suppose, if possible, that p/q is a positive fraction, in its lowest terms, such that $(p/q)^2 = 2$ or $p^2 = 2q^2$. It is easy to see that then we must have $(2q-p)^2 = 2(p-q)^2$, and so $(2q-p)/(p-q)$ is another fraction having the same property. But clearly $q < p < 2q$, and so $p - q < q$. Hence we obtain another fraction equal to p/q and having a *smaller* denominator, which contradicts the assumption that p/q is in its lowest terms.

DEFINITION. *Any point P on the line L which is not a rational point is called an* **irrational point.** *The length A_0P is called an* **irrational number.**

Examples II. 1. Show from first principles, without assuming the general theorem proved above, that $\sqrt{2}$ is not a rational number.

2. Give a similar proof for $\sqrt[3]{2}$.

3. Prove generally that, if p/q is a rational fraction in its lowest terms, $\sqrt[3]{(p/q)}$ cannot be rational unless p and q are both perfect cubes.

4. The square root of an integer must be either integral or irrational (i.e. it cannot be a *rational fraction*).

[For suppose, if possible, $\sqrt{n}=p/q$, where p, q are positive integers without a common factor. Then $nq^2=p^2$. Hence p^2 divides n, as p and q have no common factor; i.e. $n=\lambda p^2$, where λ is an integer, and so $\lambda q^2=1$, which shows that $\lambda=1$, $q=1$, $n=p^2$, and so $\sqrt{n}=p$.]

5. A more general proposition, due to Gauss, is the following: *if*

$$x^n+p_1x^{n-1}+p_2x^{n-2}+\ldots+p_n=0$$

is any algebraical equation with integral coefficients, it cannot have a rational but not integral root.

[For suppose that the equation has a root a/b, where a and b are integers without a common factor, and b is positive. Writing a/b for x, and multiplying by b^{n-1}, we obtain

$$-\frac{a^n}{b}=p_1a^{n-1}+p_2a^{n-2}b+\ldots+p_nb^{n-1},$$

a fraction in its lowest terms equal to an integer, which is absurd. Thus $b=1$, and the root is a. It is evident that a must be a divisor of p_n.]

6. Show that if $p_n=1$ and $p_1+p_2+\ldots+p_n \neq -1$, the equation cannot have a rational root.

7. Find the rational roots (if any) of

$$x^4-4x^3-8x^2+13x+10=0.$$

[The roots can only be integral, and so ± 1, ± 2, ± 5, ± 10 are the only possibilities: whether these are roots can be determined by trial. It is clear that we can in this way determine the rational roots of any such equation.]

3. Quadratic surds. If a is any rational number, the two numbers $\pm\sqrt{a}$ are either rational or irrational, and (as appears from what precedes) *generally* the latter. Numbers of this kind, when irrational, are called *pure quadratic surds*. A number $a\pm\sqrt{b}$, the sum of a rational number and a pure quadratic surd, is sometimes called a *mixed quadratic surd*.

The only kind of irrationals for whose existence the geometrical arguments of the preceding section have given us any warrant are these quadratic surds, pure and mixed, and the

more complicated irrationals which may be expressed in a form involving the repeated extraction of square roots, such as

$$\sqrt{2}+\sqrt{2+\sqrt{2}}+\sqrt{2+\sqrt{2+\sqrt{2}}}.$$

It is easy to construct geometrically a line whose length is equal to any number of this form, as the reader will easily see for himself. That *only* irrationals of these kinds can be constructed by Euclidean methods (i.e. by geometrical constructions with ruler and compasses) is a point the proof of which must be deferred for the present*. This particular property of quadratic surds naturally makes them peculiarly interesting.

Examples III. 1. Give geometrical constructions for

$$\sqrt{2},\ \sqrt{2+\sqrt{2}},\ \sqrt{2+\sqrt{2+\sqrt{2}}}.$$

2. The quadratic equation $ax^2+2bx+c=0$ has two real roots† if $b^2-ac>0$. Suppose a, b, c rational. Nothing is lost by taking all three to be integers, for we can multiply the equation by the L.C.M. of their denominators.

The reader will remember that the roots are $\{-b\pm\sqrt{(b^2-ac)}\}/a$. It is easy to construct these lengths geometrically, first constructing $\sqrt{(b^2-ac)}$. A much more elegant, though less straightforward, construction is the following.

Draw a circle of unit radius, a diameter PQ, and the tangents at the ends of the diameters.

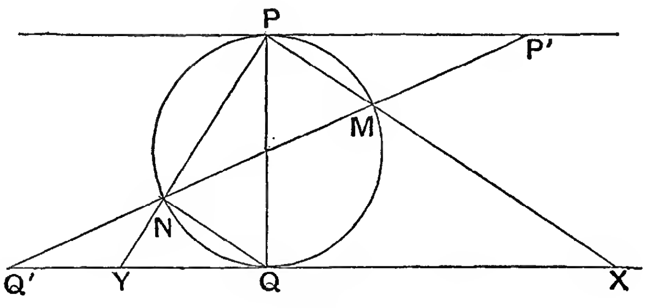


FIG. 3.

* See Chap. II, Misc. Exs. 41.

† I.e. there are two values of x for which $ax^2+2bx+c=0$. If $b^2-ac<0$ there are no such values of x. The reader will remember that in books on elementary algebra the equation is said to have two 'imaginary' roots. The meaning to be attached to this statement will be explained in Chap. III.

When $b^2=ac$ the equation has only one root. For the sake of uniformity it is generally said in this case to have 'two equal' roots, but this is a mere convention.

Take $PP' = -2a/b$ and $QQ' = -c/2b$, having regard to sign. Join $P'Q'$, cutting the circle in M and N. Draw PM and PN, cutting QQ' in X and Y. Then QX and QY are the roots of the equation with their proper signs†.*

The proof is simple and we leave it as an exercise to the reader. Another, perhaps even simpler, construction is the following. *Take a line AB of unit length. Draw $BC = -2b/a$ perpendicular to AB, and $CD = c/a$ perpendicular to BC and in the same direction as BA. On AD as diameter describe a circle cutting BC in X and Y. Then BX and BY are the roots.*

3. If ac is positive PP' and QQ' will be drawn in the same direction. Verify that if $b^2 < ac$ $P'Q'$ will not meet the circle, while if $b^2 = ac$ it will be a tangent. Verify also that if $b^2 = ac$ the circle in the second construction will touch BC.

4. Some theorems concerning quadratic surds. We shall assume that the reader is familiar with the ordinary rules for the manipulation of quadratic surds; such, e.g., as are expressed by the equations

$$\sqrt{(pq)} = \sqrt{p} \cdot \sqrt{q}, \quad \sqrt{(p^2 q)} = p\sqrt{q}.$$

He will find it a useful exercise at this stage to supply proofs of these equations.

Similar and dissimilar surds. Two pure quadratic surds are said to be *similar* if they can be expressed as rational multiples of the same surd, and otherwise *dissimilar*. Thus

$$\sqrt{8} = 2\sqrt{2}, \quad \sqrt{\tfrac{25}{2}} = \tfrac{5}{2}\sqrt{2},$$

and so $\sqrt{8}$, $\sqrt{\frac{25}{2}}$ are similar surds. On the other hand, if M and N are integers which have no common factor, and neither of which is a perfect square, $\sqrt{M}$ and $\sqrt{N}$ are dissimilar surds. For suppose, if possible,

$$\sqrt{M} = \frac{p}{q}\sqrt{\frac{\alpha}{\beta}}, \quad \sqrt{N} = \frac{r}{s}\sqrt{\frac{\alpha}{\beta}},$$

where all the letters denote integers.

Then $\sqrt{MN}$ is evidently rational, and therefore (Ex. II. 4)

* The figure is drawn to suit the case in which b and c have the same and a the opposite sign. The reader should draw figures for other cases.

† I have taken this construction from Klein's *Leçons sur certaines questions de Géométrie Élémentaire* (French translation by J. Griess, Paris, 1896).

integral. Thus $MN = P^2$ where P is an integer. Let a, b, c, ... be the prime factors of P, so that

$$MN = a^2 b^2 c^2 \ldots.$$

Then MN is divisible by a^2, and therefore either (1) M is divisible by a^2, or (2) N is divisible by a^2, or (3) M and N are both divisible by a. The last case may be ruled out, since M and N have no common factor. This argument may be applied to each of the factors a^2, b^2, c^2, ... Ultimately we see that M must be divisible by some of these factors and N by the rest. Thus

$$M = \lambda P_1^2, \quad N = \lambda P_2^2,$$

where P_1^2 denotes the product of some of the factors a^2, b^2, c^2, ... and P_2^2 the product of the rest. Since M and N have no common factor we must have $\lambda = 1$, $M = P_1^2$, $N = P_2^2$; i.e. M and N are both perfect squares, which is contrary to our hypotheses.

THEOREM. *If A, B, C, D are rational and*

$$A + \sqrt{B} = C + \sqrt{D},$$

then either (i) $A = C$, $B = D$ *or* (ii) B *and* D *are both squares of rational numbers.*

If A is not equal to C, let $A = C + x$. Then, $\sqrt{B} = x + \sqrt{D}$,

or
$$B = x^2 + D + 2x\sqrt{D};$$
i.e.
$$\sqrt{D} = (B - D - x^2)/2x,$$

which is rational, and therefore D is the square of a rational number. In this case $\sqrt{B} = C - A + \sqrt{D}$ is also rational. On the other hand, if $A = C$ it is obvious that $B = D$.

Corollaries. (i) If $A + \sqrt{B} = C + \sqrt{D}$, then $A - \sqrt{B} = C - \sqrt{D}$ (unless $\sqrt{B}$ and $\sqrt{D}$ are both rational).

(ii) The equation $\sqrt{B} = C + \sqrt{D}$ is impossible unless $C = 0$, $B = D$, or both $\sqrt{B}$ and $\sqrt{D}$ are rational.

Examples IV. 1. Prove *ab initio* that $\sqrt{2}$ and $\sqrt{3}$ are not similar surds.

2. Prove that $\sqrt{x}$ and $\sqrt{(1/x)}$ are similar surds (unless both are rational).

3. If a and b are positive and rational $\sqrt{a} + \sqrt{b}$ cannot be rational unless $\sqrt{a}$ and $\sqrt{b}$ are rational. The same is true of $\sqrt{a} - \sqrt{b}$, unless $a = b$.

4. If $$\sqrt{A}+\sqrt{B}=\sqrt{C}+\sqrt{D},$$
then either (a) $A=C$ and $B=D$, or (b) $A=D$ and $B=C$, or (c) $\sqrt{A}$, $\sqrt{B}$, $\sqrt{C}$, $\sqrt{D}$ are all rational or all similar surds.

[Square the given equation and apply the theorem above.]

5. A quadratic surd cannot be the sum of two dissimilar quadratic surds.

6. Neither $(a+\sqrt{b})^3$ nor $(a-\sqrt{b})^3$ can be rational unless $\sqrt{b}$ is rational.

7. Prove that if $x=p+\sqrt{q}$, where p and q are rational, x^m, where m is any integer, can be expressed in the form $P+Q\sqrt{q}$, where P and Q are rational. For example,
$$(p+\sqrt{q})^2=p^2+q+2p\sqrt{q}, \quad (p+\sqrt{q})^3=p^3+3pq+(3p^2+q)\sqrt{q}.$$
Deduce that any polynomial in x with rational coefficients (i.e. any expression of the form
$$a_0x^n+a_1x^{n-1}+\dots+a_n,$$
where $a_0, \dots a_n$ are rational numbers), can be expressed in the form $P+Q\sqrt{q}$.

8. Express $1/(p+\sqrt{q})$ in the same form.

$$\left[\text{We obtain } \frac{1}{p+\sqrt{q}}=\frac{p}{p^2-q}-\frac{\sqrt{q}}{p^2-q}.\right]$$

9. Deduce from Exs. 7 and 8 that any expression of the form $G(x)/H(x)$, where $G(x)$ and $H(x)$ are polynomials in x with rational coefficients, can be expressed in the form $P+Q\sqrt{q}$, where P and Q are rational.

10. If $a+\sqrt{b}$, where b is not a perfect square, is the root of an algebraical equation with rational coefficients, then $a-\sqrt{b}$ is another root of the same equation.

11. If p, q, and p^2-q are positive we can express
$$\sqrt{p+\sqrt{q}}$$
in the form $\sqrt{x}+\sqrt{y}$, where
$$x=\tfrac{1}{2}\{p+\sqrt{(p^2-q)}\}, \quad y=\tfrac{1}{2}\{p-\sqrt{(p^2-q)}\}.$$

12. Determine the conditions that it may be possible to express $\sqrt{p+\sqrt{q}}$, where p and q are rational, in the form $\sqrt{x}+\sqrt{y}$, where x and y are rational.

13. If a^2-b is positive, the necessary and sufficient conditions that
$$\sqrt{(a+\sqrt{b})}+\sqrt{(a-\sqrt{b})}$$
should be rational are that a^2-b and $\frac{1}{2}(a+\sqrt{a^2-b})$ should both be squares of rational numbers.

5. Irrational numbers in general. The arguments which led us to believe in the existence of quadratic surds, and corresponding points on the line L, were based on considerations of elementary geometry. There is, however, another way of looking at the matter which is even more instructive and important, as it leads us to consider classes of irrational numbers far more general than quadratic surds.

Consider the equation $x^2 = 2$. We have already seen that there is no rational number x which satisfies this equation. The square of any rational number is either less than or greater than 2. We can therefore divide the rational numbers into two classes, those whose squares are less than 2, and those whose squares are greater than 2. We call these two classes *the class T*, or *the lower class*, and *the class U*, or *the upper class*. It is obvious that every member of U is greater than all the members of T. Moreover, we can find a member of the class T whose square, though less than 2, differs from 2 by as little as we please. In fact, if we carry out the ordinary arithmetical process for the extraction of the square root of 2 we obtain a series of rational numbers, viz.

$$1,\ 1{\cdot}4,\ 1{\cdot}41,\ 1{\cdot}414,\ 1{\cdot}4142,\ \ldots$$

whose squares

$$1,\ 1{\cdot}96,\ 1{\cdot}9881,\ 1{\cdot}999396,\ 1{\cdot}99996164,\ \ldots$$

are all less than 2, but approach nearer and nearer to it; and by taking a sufficient number of the figures given by the process, we can obtain as close an approximation as we want. Similarly we can find a member of the class U whose square, though greater than 2, differs from 2 by as little as we please. It is sufficient to increase the last figure, in the series of approximations given above, by unity: we obtain

$$2,\ 1{\cdot}5,\ 1{\cdot}42,\ 1{\cdot}415,\ 1{\cdot}4143,\ \ldots.$$

Or again, we can find a member of T and a member of U which differ from one another by as little as we please.

This follows at once from the fact that *every* rational number belongs to one class or the other. A formal proof may be supplied as follows. Take any member x of T and any member y of U. Let n be any positive integer, and consider the numbers

$$x,\quad x+\frac{1}{n}(y-x),\quad x+\frac{2}{n}(y-x),\quad \ldots\ldots,\quad x+\frac{n-1}{n}(y-x),\quad y.$$

Each of these is rational and belongs either to T or to U. Let $x+(r/n)(y-x)$ be the *first* which belongs to T. Then $x+\{(r+1)/n\}(y-x)$ belongs to U. Thus we have found a member of T and a member of U which differ by $(y-x)/n$. And by taking a large enough value of n we can make this difference as small as we like.

We add a formal proof that an x can be found in T and a y in U such that $x^2<2$ and $y^2>2$, but both squares differ from 2 by as little as we please.

Suppose we want each difference to be less than ϵ (where ϵ may be, say, $\cdot 01$ or $\cdot 0001$ or $\cdot 00001$). We can, in virtue of what precedes, choose x and y so that

$$y - x < \tfrac{1}{4}\epsilon.$$

We may obviously suppose both x and y less than 2, since $x^2 < 2$ and y is nearly equal to x. Then

$$y^2 - x^2 = (y - x)(y + x) < 4(y - x) < \epsilon,$$

and since $x^2 < 2$ and $y^2 > 2$ it follows *à fortiori* that $2 - x^2$ and $y^2 - 2$ are each less than ϵ.

We have thus divided all the positive rational points on L into two classes T and U such that (i) the class U lies entirely to the right of the class T, (ii) we can find a pair of points, one in T and one in U, whose distance from one another is as small as we please. And our common-sense notion of the attributes of a straight line demands *the existence of a number x and a corresponding point P such that P divides the class T from the class U.*

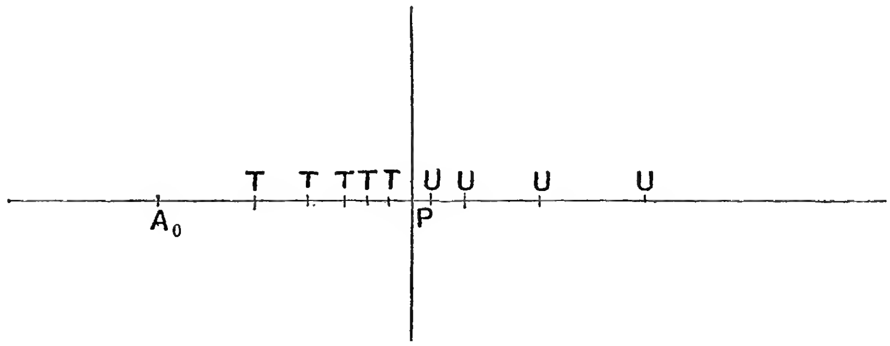


FIG. 4.

But (1) this number x cannot be rational. For if it were, P would belong either to the class T or the class U, let us say the former. Then $x^2 < 2$, or $x^2 = 2 - \delta$, say, where δ is some positive number. But we can find a member of the class T whose square is as near to 2 as we like; and therefore we can find such a member of T whose square is greater than $2 - \delta$, i.e. greater than x^2. That is to say, we can find a member of T which *lies to the right of P*: which is absurd. Hence P cannot belong to T. Similarly it cannot belong to U.

Again (2) x^2 cannot be either less than or greater than 2. If it were less than 2 we could, as above, find members of T to the right of P. This hypothesis is therefore untenable, and x^2 is not less than 2. Similarly it is not greater than 2.

Hence *there is a point P and a number x such that*

$$A_0P = x, \quad x^2 = 2.$$

This number x we denote by $\sqrt{2}$.

Examples V. 1. Find the difference between 2 and the squares of the decimals given in § 5 as approximations to $\sqrt{2}$.

2. Find the differences between 2 and the squares of

$$\tfrac{1}{1}, \tfrac{3}{2}, \tfrac{7}{5}, \tfrac{17}{12}, \tfrac{41}{29}, \tfrac{99}{70}.$$

3. Show that if m/n is a good approximation to $\sqrt{2}$, then $(m+2n)/(m+n)$ is a better one, and that the errors in the two cases are in opposite directions. Apply this result to continue the series of approximations in the last example.

4. If x and y are approximations to $\sqrt{2}$, by defect and by excess respectively, and $2-x^2 < \epsilon$, $y^2-2 < \epsilon$, then $y-x < \epsilon$.

5. The equation $x^2=4$ is satisfied by $x=2$. Examine how far the argument of the preceding sections applies to this equation (writing 4 for 2 throughout).

[We define the classes T, U as above. But in this case they do not include *all* rationals. The rational number 2 is an exception, since 2^2 is neither less than or greater than 4. And as before we are led to suppose the existence of a *dividing point*. But we cannot, of course, prove that this is not a rational point. It is, in fact, the point $x=2$.]

6. But the preceding argument may be applied to equations other than $x^2 = 2$, almost word for word; for example to $x^2 = N$, where N is any integer which is not a perfect square, or to

$$x^3 = 3, \quad x^3 = 7, \quad x^4 = 23,$$

or, as we shall see later on, to $x^3 = 3x + 8$. We are thus led to believe in the existence of points P on L such that $x = A_0P$ satisfies equations such as these, even when these lengths cannot be constructed by means of elementary geometrical methods.

The reader will no doubt remember that in treatises on elementary algebra the root of such an equation as $x^q = n$ is denoted by $\sqrt[q]{n}$ or $n^{1/q}$, and that a meaning is attached to such symbols as

$$n^{p/q}, \quad n^{-p/q}$$

by means of the equations

$$n^{p/q} = \sqrt[q]{n^p}, \quad n^{p/q}\, n^{-p/q} = 1.$$

And he will remember how, in virtue of these definitions, the 'laws of indices' such as

$$n^r \times n^s = n^{r+s}, \quad (n^r)^s = n^{rs}$$

are extended to cover the case in which r and s are any rational numbers whatever.

7. The continuum. The aggregate of points contained in a straight line L is called a **linear continuum.** It contains (1) the rational points, (2) the irrational points for whose existence we have the evidence summarized in the preceding sections, and the corresponding negative irrational points on the left of A_0, (3) all other points of the line, if any such there be. To each of these points corresponds a length measured from A_0, and capable of numerical measurement in terms of our unit-length A_0A_1. The measures of these lengths are the **real numbers,** positive or negative, integral, rational or irrational. The aggregate of all these numbers is called an **arithmetical continuum.** All the numbers contained in this arithmetical continuum may be operated with according to the ordinary rules of elementary algebra.

The substance of the preceding sections is not intended as a complete or rigorous analysis of the nature of either the linear or the arithmetical continuum. Such an analysis would be altogether beyond the scope of this book. What has been said is intended simply to remind the reader of some of the ideas on the subject which he no doubt already possesses, and to attempt to make them, and some of the obvious consequences which are involved in them, more explicitly present to his mind.

In order to show the incompleteness of the analysis of the numbers of the arithmetic continuum which has been given, we need only consider a few examples.

(i) Let us consider a more complicated surd expression, such as

$$z = \sqrt[3]{(4+\sqrt{15})} + \sqrt[3]{(4-\sqrt{15})}.$$

Our argument for supposing that the expression for z has a meaning, and that a point P exists on the line such that $A_0P = z$, might be as follows. We first show, as above, that there is a point P_1 such that if $y = A_0P_1$, $y^2 = 15$, and we can then determine points corresponding to the numbers $4+\sqrt{15}$, $4-\sqrt{15}$. Now consider the equation in z_1

$$z_1^3 = 4+\sqrt{15}.$$

The right-hand side of this equation is not rational: but exactly the same reasoning which leads us to suppose that the line contains a point x for which $x^3=2$ (or any other rational number) also leads us to the conclusion that it contains a point z_1 for which $z_1{}^3=4+\surd 15$. Thus we find a point P' such that

$$z_1=A_0P'=\sqrt[3]{(4-\surd 15)}.$$

Similarly we find a point P'' such that

$$z_2=A_0P''=\sqrt[3]{(4-\surd 15)},$$

and taking $A_0P=A_0P'+A_0P''$ we have finally

$$A_0P=z=\sqrt[3]{(4+\surd 15)}+\sqrt[3]{(4-\surd 15)}.$$

Now it is easy to verify that

$$z^3=3z+8.$$

And we might have given a direct proof of the existence of a unique number z such that $z^3=3z+8$. It is easy to see that there cannot be two such numbers. For if $z_1{}^3=3z_1+8$ and $z_2{}^3=3z_2+8$, we find on subtracting and dividing by z_1-z_2 that $z_1{}^2+z_1z_2+z_2{}^2=3$. But if z_1 and z_2 are positive $z_1{}^3>8$, $z_2{}^3>8$ and therefore $z_1>2$, $z_2>2$, $z_1{}^2+z_1z_2+z_2{}^2>12$, and so the equation just found is impossible. And it is easy to see that neither z_1 nor z_2 can be negative. For if z_1 is negative and equal to $-\zeta$, ζ is positive and $\zeta^3-3\zeta+8=0$, or $3-\zeta^2=8/\zeta$. Hence $3-\zeta^2>0$, and so $\zeta<2$. But then $8/\zeta>4$ and cannot be equal to $3-\zeta^2$, which is less than 3.

Hence there is at most *one* z such that $z^3=3z+8$. And it cannot be rational. For any rational root of this equation must be integral and a factor of 8 (Ex. II. 5), and it is easy to verify that no one of ± 1, ± 2, ± 4, ± 8, is a root.

Thus $z^3=3z+8$ has at most one root and that root is not rational. We can now define the positive rational numbers x into two classes T, U according as $x^3<3x+8$, or $x^3>3x+8$. It is easy to see that if $x^3>3x+8$ and y is any number greater than x, then also $y^3>3y+8$. For suppose if possible $y^3\leqq 3y+8$. Then since $x^3>3x+8$ we obtain on subtracting $y^3-x^3<3(y-x)$, or $y^2+xy+x^2<3$, which is impossible, since y is positive and $x>2$ (since $x^3>8$). Similarly we can show that if $x^3<3x+8$ and $y<x$ then also $y^3<3y+8$.

Thus we have separated the rational numbers into two classes similar to the classes T, U of § 5. And we conclude, as there, that there is a number z which is greater than any number of T, and less than any number of U, and which satisfies the equation $z^3=3z+8$.

The reader who knows how to solve cubic equations by Cardan's method will be able to obtain directly from the equation the expression

$$z=\sqrt[3]{(4+\surd 15)}+\sqrt[3]{(4-\surd 15)}.$$

(ii) The direct argument applied above to the equation $x^3 = 3x + 8$ could be applied (though the application would be a little more difficult) to the equation

$$x^5 = x + 16,$$

and would lead us to the conclusion that a unique positive number exists which satisfies this equation. In this case, however, it is not possible to obtain a simple explicit expression for x composed of any combination of surds. It can in fact be proved (though the proof is difficult) that it is *generally* impossible to find such an expression for the root of an equation of higher degree than 4.

Thus, besides irrational numbers which can be expressed as pure or mixed quadratic or other surds, or combinations of such surds, there are others which cannot be so expressed. It is *only in very special cases* that such expressions can be found.

(iii) But even when we have added to our list of irrational numbers roots of equations (such as $x^5 = x + 16$) which cannot be explicitly expressed as surds, we have not exhausted the different kinds of irrational numbers contained in the continuum. Let us draw a circle whose diameter is equal to A_0A_1, i.e. to unity. It is natural to suppose that the circumference of such a circle has a length capable of numerical measurement as much as the diagonal of a square described on A_0A_1. This length is usually denoted by π. And it has been shown (though the proof is unfortunately long and difficult) that this number π is not the root of any algebraical equation with integral coefficients, so that we cannot have, for example, any such equation as

$$\pi^2 = n, \quad \pi^3 = n, \quad \pi^5 = \pi + n,$$

where n is an integer. If we take a point P such that $A_0P = \pi$, we have found a point which is not rational nor yet belongs to any of the classes of irrationals which we have so far considered. And this number π is no isolated or exceptional case. Any number of other examples can be constructed. In fact it is only special classes of irrational numbers which are roots of equations of this kind, just as it is only a still smaller class which can be expressed by means of surds.

Examples VI. 1. Show that $x = \sqrt[3]{(5+2\sqrt{6})} + \sqrt[3]{(5-2\sqrt{6})}$ satisfies the equation

$$x^3 = 3x + 10,$$

and apply to this equation arguments similar to those used in § 7 (i). And, more generally, $x = \sqrt[3]{\{n + \sqrt{(n^2-1)}\}} + \sqrt[3]{\{n - \sqrt{(n^2-1)}\}}$ satisfies

$$x^3 = 3x + 2n.$$

Consider this equation similarly, n being any positive integer.

2. Consider the equation

$$x^2 - 4x + 3 = 0.$$

It is easy to see that $x=1$ and $x=3$ are roots of this equation. If we divide the rational numbers into two classes T and U according as

$$x^2 - 4x + 3 \lesseqgtr 0,$$

we see that T contains all rational numbers between 1 and 3 and U all less than 1 or greater than 3. In this case we are not led to any irrational number, the numbers which divide the classes being 1 and 3. But if we consider instead the equation

$$x^2 - 4x + 1 = 0,$$

(of which the roots are $2 \pm \sqrt{3}$), we again have two points of division, in this case each irrational: and we might argue directly from the equation to the existence of two such numbers by dividing up the rational numbers into classes T, U as above.

8. **The continuous real variable.** The 'real numbers' may be regarded from two points of view. We may think of them *as an aggregate*, the 'arithmetical continuum' defined in the preceding section, or *individually*. And when we think of them individually, we may think either of a particular *specified* number (such as 1, $-\frac{1}{2}$, $\sqrt{2}$, or π) or we may think of *any* number, *an unspecified* number, *the number* x. This last is our point of view when we make such assertions as 'x is a number,' 'x is the measure of a length,' 'x may be rational or irrational.' The x which occurs in propositions such as these is called *the continuous real variable*: and the individual numbers are called the *values* of the variable.

A 'variable,' however, need not necessarily be continuous. Instead of considering the aggregate of *all* real numbers, we might consider some partial aggregate contained in the former aggregate, such as the aggregate of rational numbers, or the aggregate of positive integers. Let us take the last case. Then in statements about *any* positive integer, or *an unspecified* positive

integer, such as 'n is either odd or even,' n is called the variable, a *positive integral variable*, and the individual positive integers are its values.

In fact, this x and n are only examples of variables, the variable whose 'field of variation' is formed by all the real numbers, and that whose field is formed by the positive integers. These are the most important examples, but we have often to consider other cases. In the theory of decimals, for instance, we may denote by x any figure in the expression of any number as a decimal. Then x is a variable, but a variable which has only ten different values, viz. 0, 1, 2, 3, 4, 5, 6, 7, 8, 9. The reader should think of other examples of variables with different fields of variation. He will find interesting examples in ordinary life. For instance—policeman x, the driver of cab x, star x in Herschel's catalogue, the year x, the xth day of the week.

MISCELLANEOUS EXAMPLES ON CHAPTER I.

1. If $a, b, c, \dots k$ and $A, B, C, \dots K$ are two sets of numbers, and all of the first set are positive, then

$$\frac{aA+bB+\dots+kK}{a+b+\dots+k}$$

lies between the algebraically least and greatest of $A, B, \dots, K$.

2. What are the conditions that $ax+by+cz=0$, (1) for all values of x, y, z; (2) for all values of x, y, z subject to $\alpha x+\beta y+\gamma z=0$; (3) for all values of x, y, z subject to both $\alpha x+\beta y+\gamma z=0$ and $Ax+By+Cz=0$?

3. Any positive rational number can be expressed in one and only one way in the form

$$a_1+\frac{a_2}{1.2}+\frac{a_3}{1.2.3}+\dots+\frac{a_k}{1.2.3\dots k},$$

where $a_1, a_2, \dots, a_k$ are integers, and

$$0 \leqq a_1, \quad 0 \leqq a_2 < 2, \quad 0 \leqq a_3 < 3, \dots 0 \leqq a_k < k.$$

[For example $\dfrac{47}{21}=\dfrac{11280}{7\,!}=\dfrac{7\cdot 1611+3}{7\,!}=\dfrac{3}{7\,!}+\dfrac{6\cdot 268+3}{6\,!}$

$$=\frac{3}{7\,!}+\frac{3}{6\,!}+\frac{5\cdot 53+3}{5\,!}=\frac{3}{7\,!}+\frac{3}{6\,!}+\frac{3}{5\,!}+\frac{4\cdot 13+1}{4\,!}$$

$$=\frac{3}{7\,!}+\frac{3}{6\,!}+\frac{3}{5\,!}+\frac{1}{4\,!}+\frac{1}{3\,!}+\frac{0}{2\,!}+2,$$

by continuing the same process. It is evident that k is *at most* equal to the largest prime factor of the denominator of the number given.]

4. Any positive rational number can be expressed in one and one way only as a simple continued fraction

$$a_1+\frac{1}{a_2+}\frac{1}{a_3+}\dots\frac{1}{+a_n},$$

where a_1, a_2, ... are positive integers, of which the first only may be zero.

[Accounts of the theory of such continued fractions will be found in text-books of algebra.]

5. Find the rational roots (if any) of $9x^3-6x^2+15x-10=0$.

[Put $3x=y$ and apply the method of Ex. II. 7 to the resulting equation in y.]

6. A line AB is divided at C *in aurea sectione* (Euc. II. 11)—i.e. so that $AB\,.\,AC=BC^2$. Show that the ratio AC/AB is irrational.

[A direct geometrical proof will be found in Bromwich's *Infinite Series*, § 143, p. 363.]

7. A is irrational. In what circumstances can $\dfrac{aA+b}{cA+d}$, where a, b, c, d are rational, be rational?

8. Express $\surd p$, $\surd q$ in the form $ax+(b/x)$, where a, b are rational, and $x=\surd p+\surd q$.

9. If $\surd p$, $\surd q$ are dissimilar surds, and $a+b\surd p+c\surd q+d\surd(pq)=0$, where a, b, c, d are rational, then $a=0$, $b=0$, $c=0$, $d=0$.

[Express $\surd p$ in the form $M+N\surd q$, where M and N are rational, and apply the theorem of § 4.]

10. Show that if $a\surd 2+b\surd 3+c\surd 5=0$, where a, b, c are rational numbers, then $a=0$, $b=0$, $c=0$.

11. Any polynomial in $\surd p$ and $\surd q$, with rational coefficients, (i.e. any sum of a finite number of terms of the form $A(\surd p)^m(\surd q)^n$, where m and n are integers, and A rational) can be expressed in the form

$$a+b\surd p+c\surd q+d\surd pq,$$

where a, b, c, d are rational.

12. Express $\dfrac{a+b\surd p+c\surd q}{d+e\surd p+f\surd q}$, where a, b, etc. are rational, in the form

$$A+B\surd p+C\surd q+D\surd pq,$$

where A, B, C, D are rational.

[Evidently

$$\frac{a+b\surd p+c\surd q}{d+e\surd p+f\surd q}=\frac{(a+b\surd p+c\surd q)(d+e\surd p-f\surd q)}{(d+e\surd p)^2-f^2q}=\frac{\alpha+\beta\surd p+\gamma\surd q+\delta\surd pq}{\epsilon+\zeta\surd p},$$

where α, β, etc. are rational numbers which can easily be found. The required reduction may now be easily completed by multiplication of numerator and denominator by $\epsilon-\zeta\surd p$.]

For example, prove that

$$\frac{1}{1+\sqrt{2}+\sqrt{3}}=\frac{1}{2}+\frac{1}{4}\sqrt{2}-\frac{1}{4}\sqrt{6}.$$

13. If a, b, x, y are rational numbers such that

$$(ay-bx)^2+4(a-x)(b-y)=0,$$

prove that either (i) $x=a$, $y=b$, or (ii) $1-ab$ and $1-xy$ are squares of rational numbers. (*Math. Tripos*, 1903.)

[If we write $a-x=\xi$, $b-y=\eta$, we obtain

$$a^2\eta^2+b^2\xi^2+(4-2ab)\xi\eta=0.$$

Solving this equation for the ratio ξ/η we find that ξ/η (which we know to be rational) involves the quantity

$$\sqrt{\{(2-ab)^2-a^2b^2\}}=2\sqrt{(1-ab)}.$$

Hence $1-ab$ must be the square of a rational quantity. The only alternative is $\xi=\eta=0$.

But the equation given may also be written in the form

$$x^2\eta^2+y^2\xi^2+(4-2xy)\xi\eta=0.$$

Hence we deduce the same conclusion for $\sqrt{(1-xy)}$.]

14. If all the values of x and y given by

$$ax^2+2hxy+by^2=1, \quad a'x^2+2h'xy+b'y^2=1,$$

(where a, h, b, a', h', b' are rational) are rational, then

$$(h-h')^2-(a-a')(b-b'), \quad (ab'-a'b)^2+4(ah'-a'h)(bh'-b'h),$$

are both squares of rational quantities. (*Math. Tripos*, 1899.)

15. Show that $\sqrt{2}$ and $\sqrt{3}$ are cubic functions of $\sqrt{2}+\sqrt{3}$, with rational coefficients, and that $\sqrt{2}-\sqrt{6}+3$ is the ratio of two linear functions of $\sqrt{2}+\sqrt{3}$. (*Math. Tripos*, 1905.)

16. The expression

$$\sqrt{a+2m\sqrt{a-m^2}}+\sqrt{a-2m\sqrt{a-m^2}}$$

is equal to $2m$ if $2m^2>a>m^2$, and to $2\sqrt{(a-m^2)}$ if $a>2m^2$.

17. Show that any polynomial in $\sqrt[3]{2}$, with rational coefficients, can be expressed in the form

$$a+b\sqrt[3]{2}+c\sqrt[3]{4},$$

where a, b, c are rational.

More generally, any polynomial in $\sqrt[m]{p}$, with rational coefficients, can be expressed in the form

$$a_0+a_1x+a_2x^2+\ldots+a_{m-1}x^{m-1},$$

where a_0, a_1, ... are rational and $x=\sqrt[m]{p}$. For any such polynomial is of the form

$$b_0+b_1x+b_2x^2+\ldots+b_kx^k,$$

where the b's are rational. If $k \leqq m-1$ this is already of the form required. If $k > m-1$, let x^r be any power of x higher than the $(m-1)$th. Then $r = \lambda m + s$, where λ is an integer and $0 \leqq s \leqq m-1$: and $x^r = x^{\lambda m+s} = p^\lambda x^s$. Hence we can get rid of all powers of x higher than the $(m-1)$th.

18. Express $(\sqrt[3]{2}-1)^5$ in the form $a + b\sqrt[3]{2} + c\sqrt[3]{4}$, where a, b, c are rational.

19. Express $(\sqrt[3]{2}-1)/(\sqrt[3]{2}+1)$ in the same form.

[Multiply numerator and denominator by $\sqrt[3]{4} - \sqrt[3]{2} + 1$.]

20. If
$$a + b\sqrt[3]{2} + c\sqrt[3]{4} = 0,$$
where a, b, c are rational, then $a=0$, $b=0$, $c=0$.

[Let $y = \sqrt[3]{2}$. Then $y^3 = 2$ and
$$cy^2 + by + a = 0.$$
Hence $2cy^2 + 2by + ay^3 = 0$ or
$$ay^2 + 2cy + 2b = 0.$$

Multiplying these two quadratic equations by a and c and subtracting we obtain $(ab - 2c^2)y + a^2 - 2bc = 0$, or $y = -(a^2 - 2bc)/(ab - 2c^2)$, a rational number, which is impossible. The only alternative is that $ab - 2c^2 = 0$, $a^2 - 2bc = 0$.

Hence $ab = 2c^2$, $a^4 = 4b^2c^2$. If $ab \neq 0$ we can divide the second equation by the first, which gives $a^3 = 2b^3$: and this is impossible, since $\sqrt[3]{2}$ cannot be equal to the rational quantity a/b. Hence $ab = 0$, $c = 0$, and it follows from the original equation that a, b, and c are all zero.

As a corollary, if $a + b\sqrt[3]{2} + c\sqrt[3]{4} = d + e\sqrt[3]{2} + f\sqrt[3]{4}$, then $a = d$, $b = e$, $c = f$.

It may be proved, more generally, that if
$$a_0 + a_1 p^{1/m} + \dots + a_{m-1} p^{(m-1)/m} = 0,$$
p not being a perfect mth power, then $a_0 = a_1 = \dots = a_{m-1} = 0$; but the proof is by no means so simple.]

21. Prove the theorem of § 4 by the method employed in Ex. 20.

22. If $A + \sqrt[3]{B} = C + \sqrt[3]{D}$, then either $A = C$, $B = D$, or B and D are both cubes of rational quantities.

[Assume $A = C + x$, cube, and apply the result of Ex. 20.]

23. If $\sqrt[3]{A} + \sqrt[3]{B} + \sqrt[3]{C} = 0$, then either one of A, B, C is zero, and the other two equal and opposite, or $\sqrt[3]{A}$, $\sqrt[3]{B}$, $\sqrt[3]{C}$ are rational multiples of the same surd $\sqrt[3]{X}$.

24. Find rational numbers α, β such that
$$\sqrt[3]{(7 + 5\sqrt{2})} = \alpha + \beta\sqrt{2}.$$

25. If
$$(a-b^3)\,b>0,$$
$$\sqrt[3]{a+\frac{9b^3+a}{3b}\sqrt{\left(\frac{a-b^3}{3b}\right)}}+\sqrt[3]{a-\frac{9b^3+a}{3b}\sqrt{\left(\frac{a-b^3}{3b}\right)}}$$
is rational.

[Each of the quantities under a cube root can be expressed in the form
$$\left\{\alpha+\beta\sqrt{\left(\frac{a-b^3}{3b}\right)}\right\}^3,$$
where α and β are rational.]

26. If $\alpha=\sqrt[n]{A}$, any polynomial in α is the root of an equation of degree n, with rational coefficients.

[We can express the polynomial (x say) in the form
$$x=l_1+m_1\alpha+\ldots+r_1\alpha^{(n-1)/n},$$
where l_1, m_1, ... are rational.

Similarly $$x^2=l_2+m_2\alpha+\ldots+r_1\alpha^{(n-1)/n},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$x^n=l_n+m_n\alpha+\ldots+r_n\alpha^{(n-1)/n}.$$

Hence $$L_1x+L_2x^2+\ldots+L_nx^n=\Delta,$$
where Δ is the determinant
$$\begin{vmatrix} l_1 & m_1 & \ldots & r_1 \\ l_2 & m_2 & \ldots & r_2 \\ \ldots & \ldots & \ldots & \ldots \\ l_n & m_n & \ldots & r_n \end{vmatrix}$$
and L_1, L_2, ... the minors of l_1, l_2 ....]

27. Apply this process to $x=p+\sqrt{q}$. [The result is $x^2-2px+(p^2-q)=0$.]

28. Deduce from the result of the last example that if $p+\sqrt{q}=r+\sqrt{s}$, either $p=r$, $q=s$ or q, s are squares of rational quantities.

[It is easy to see that if $p+\sqrt{q}$ is not rational we must have
$$x^2-2px+(p^2-q)\equiv x^2-2rx+(r^2-s).]$$

29. Show that $y=a+bp^{1/3}+cp^{2/3}$ satisfies the equation
$$y^3-3ay^2+3y\,(a^2-bcp)-a^3-b^3p-c^3p^2+3abcp=0.$$

30. **Algebraical numbers.** We have seen that some irrational numbers (such as $\sqrt{2}$) are roots of equations of the type
$$a_0x^n+a_1x^{n-1}+\ldots+a_n=0,$$
where a_0, a_1, ..., a_n are integers. Such irrational numbers are called *algebraical* numbers: all other irrational numbers, such as π (§ 7), are called transcendental numbers. Show that if x is an algebraical number so are kx, where k is any rational number, and $x^{m/n}$, where m and n are any integers.

31. If x and y are algebraical numbers, so are $x+y$, $x-y$, xy and x/y.

[We have equations
$$a_0x^m+a_1x^{m-1}+\dots+a_m=0,$$
$$b_0y^n+b_1y^{n-1}+\dots+b_n=0,$$
where the a's and b's are integers. Write $x+y=z$, $y=z-x$ in the second, and eliminate x. We thus get an equation of similar form
$$c_0z^p+c_1z^{p-1}+\dots+c_p=0,$$
satisfied by z. Similarly for the other cases.]

32. If
$$a_0x^n+a_1x^{n-1}+\dots+a_n=0$$
where $a_0, a_1, \dots, a_n$ are any algebraical numbers, then x is an algebraical number.

[We have $n+1$ equations of the type
$$a_{0,r}\,a_r^{m_r}+a_{1,r}\,a_r^{m_r-1}+\dots+a_{m_r,r}=0,$$
($r=0, 1, \dots, n$), in which the coefficients $a_{0,r}$, $a_{1,r}$, ... are integers. Eliminate $a_0, a_1, \dots, a_n$ between these and the original equation for x.]

33. Apply this process to the equation $x^2-2x\sqrt{2}+\sqrt{3}=0$.

[The result is $x^8-16x^6+58x^4-48x^2+9=0$.]

34. Find equations, with rational coefficients, satisfied by
$$\sqrt{2}+\sqrt{3},\quad 1+\sqrt{2}+\sqrt{3},\quad \frac{\sqrt{3}+\sqrt{2}}{\sqrt{3}-\sqrt{2}},\quad \sqrt{(2+\sqrt{2})}+\sqrt{(2-\sqrt{2})},$$
$$\sqrt{\{\sqrt{3}+\sqrt{2}\}}+\sqrt{\{\sqrt{3}-\sqrt{2}\}},\quad \sqrt[3]{2}+\sqrt[3]{3},\quad 1+\sqrt[3]{2}+\sqrt[3]{3},\quad \frac{\sqrt[3]{3}+\sqrt[3]{2}}{\sqrt[3]{3}-\sqrt[3]{2}}.$$

35. If $x^3=x+1$, then $x^{3n}=a_nx+b_n+c_n/x$, where
$$a_{n+1}=a_n+b_n,\quad b_{n+1}=a_n+b_n+c_n,\quad c_{n+1}=a_n+c_n.$$

36. If $x^6+x^5-2x^4-x^3+x^2+1=0$ and $y=x^4-x^2+x-1$, then y satisfies a quadratic equation with rational coefficients. (*Math. Tripos*, 1903.)

[It will be found that $y^2+y+1=0$.]

CHAPTER II.

FUNCTIONS OF REAL VARIABLES.

9. The idea of a function. Suppose that x and y are two continuous real variables, which we may suppose to be represented geometrically by distances $A_0P = x$, $B_0Q = y$ measured from fixed points A_0, B_0 along two straight lines L, M. And

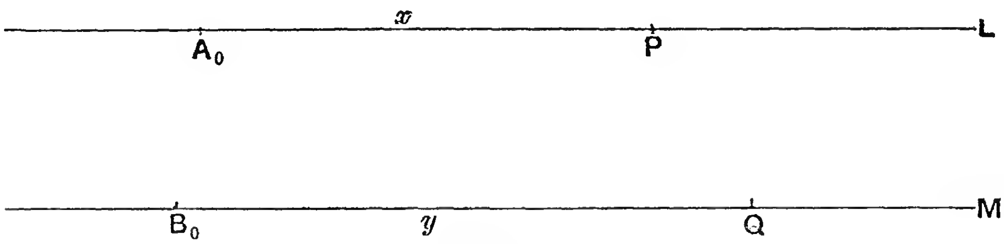


FIG. 5.

let us suppose that the positions of the points P and Q are not independent, but connected by a relation which we can imagine to be expressed as a relation between x and y: so that, when P and x are known, Q and y are also known. We might, for example, suppose that $y = x$, or $y = 2x$, or $\frac{1}{2}x$, or $x^2 + 1$. In all of these cases the value of x determines that of y. Or again, we might suppose that the relation between x and y is given, not by means of an explicit formula for y in terms of x, but by means of a geometrical construction which enables us to determine Q when P is known.

In these circumstances y is said to be a *function* of x. This notion of functional dependence of one variable upon another is perhaps the most important in the whole range of higher mathematics. In order to enable the reader to be certain that he understands it clearly we shall, in this chapter, illustrate it by means of a large number of examples.

But before we proceed to do this, we must point out that the simple examples of functions mentioned above possess three characteristics which are by no means involved in the general idea of a function, viz.:

(1) y is determined *for every value of* x;

(2) to each value of x for which y is given corresponds *one and only one value of* y;

(3) the relation between x and y is expressed by means of *an analytical formula.*

It is indeed the case that these particular characteristics are possessed by *many of the most important* functions. But the consideration of the following examples will make it clear that they are by no means essential to a function. All that is essential is that there should be some relation between x and y such that to some values of x at any rate correspond values of y.

Examples VII. 1. Let $y=x$ or $2x$ or $\frac{1}{2}x$ or x^2+1. Nothing further need be said at present about cases such as these.

2. Let $y=0$ *whatever be the value of* x. Then y is a function of x, for we can give x any value, and the corresponding value of y (viz. 0) is known. In this case the functional relation makes the same value of y correspond to all values of x. The same would be true were y equal to 1 or $-\frac{1}{2}$ or $\sqrt{2}$ instead of 0. Such a function of x is called *a constant.*

3. Let $y^2=x$. Then if x is positive this equation defines *two* values of y corresponding to each value of x, viz. $\pm\sqrt{x}$. If $x=0$, $y=0$. Hence to the particular value 0 of x corresponds *one* and only one value of y. But if x is negative there is *no* value of y which satisfies the equation. That is to say, *the function y is not defined for negative values of x.*

This function therefore possesses the characteristic (3), but not (1) or (2).

4. Consider a volume of gas maintained at a constant temperature and contained in a cylinder closed by a sliding piston*.

Let A be the area of the cross section of the piston and W its weight. The gas, held in a state of compression by the piston, exerts a certain pressure p_0 per unit of area on the piston, which balances the weight W, so that

$$W=Ap_0.$$

Let v_0 be the volume of the gas when the system is thus in equilibrium. If additional weight is placed upon the piston the latter is forced downwards. The volume (v) of the gas diminishes; the pressure (p) which it exerts

* I borrow this instructive example from Prof. H. S. Carslaw's *Introduction to the Calculus.*

upon unit area of the piston increases. Boyle's experimental law asserts that the product of p and v is very nearly constant, a correspondence which, if exact, would be represented by an equation of the type

$$pv = a \quad \text{.....................................(i)},$$

where a is a number which can be determined approximately by experiment.

Boyle's law, however, only gives a reasonable approximation to the facts provided the gas is not compressed too much. When v is decreased and p increased beyond a certain point the relation between them is no longer expressed with tolerable exactness by the equation (i). It is known that a much better approximation to the true relation can then be found by means of what is known as 'van der Waals' law,' expressed by the equation

$$\left(p + \frac{a}{v^2}\right)(v - \beta) = \gamma \quad \text{..............................(ii)},$$

where a, β, γ are numbers which can also be determined approximately by experiment.

Of course the two equations, even taken together, do not give anything like a complete account of the relation between p and v. This relation is no doubt in reality much more complicated, and its form changes, as v varies, from a form nearly equivalent to (i) to a form nearly equivalent to (ii). But, from a mathematical point of view, there is nothing to prevent us from contemplating an ideal state of things in which, for all values of v above a certain limit, V say, (i) would be exactly true, and (ii) exactly true for all values of v less than V. And then we might regard the two equations as together defining p as a function of v. It is an example of a function which *for some values of v is defined by one formula and for other values of v is defined by another.*

This function possesses the characteristic (2): to any value of v only one value of p corresponds: but it does not possess (1). For p is at any rate not defined as a function of v for negative values of v; a *negative volume* means nothing, and so negative values of v do not present themselves for consideration at all.

5. Suppose that a perfectly elastic ball is dropped (without rotation) from a height $\frac{1}{2}g\tau^2$ on to a fixed horizontal plane, and rebounds continually.

The ordinary formulae of elementary dynamics, with which the reader is probably familiar, show that

$$h = \tfrac{1}{2}gt^2$$

if $0 \leqq t \leqq \tau$,

$$h = \tfrac{1}{2}g(2\tau - t)^2$$

if $\tau \leqq t \leqq 3\tau$, and generally

$$h = \tfrac{1}{2}g(2n\tau - t)^2$$

if $(2n-1)\tau \leqq t \leqq (2n+1)\tau$, h being the depth of the ball, at time t, below its original position. Obviously h is a function of t which is only defined for positive values of t.

The reader should construct other examples of functions which occur in physical problems.

6. Suppose that y is defined as being *the largest prime factor of* x. This is an instance of a definition which only applies to a particular class of values of x, viz., *integral* values. 'The largest prime factor of $\frac{11}{3}$ or of $\sqrt{2}$ or of π' means nothing, and so our defining relation fails to define for such values of x as these. Thus this function does not possess the characteristic (1). It does possess (2), but not (3), as there is no simple formula which expresses y in terms of x.

7. Let y be defined as *the denominator of x when x is expressed in its lowest terms.* This is an example of a function which is defined if and only if x is *rational*. Thus $y=7$ if $x=-11/7$: but y is not defined for $x=\sqrt{2}$, 'the denominator of $\sqrt{2}$' being a meaningless expression.

8. Let y be defined as *the height in inches of policeman C. x, in the Metropolitan Police, at* 5.30 *p.m. on* 8 *Aug.* 1907. Then y is defined for a certain number of integral values of x, viz., $1, 2, \ldots, N$, where N is the total number of policemen in division C at that particular moment of time.

10. The graphical representation of functions. Coordinate geometry of two dimensions. Suppose that the variable y is a function of the variable x. It will generally be open to us also to regard x as a function of y, in virtue of the functional relation between x and y. But for the present we shall look at this relation from the first point of view. We shall then call x the *independent variable* and y the *dependent variable*; and, when the particular form of the functional relation is not specified, we shall express it by the general form of equation

$$y=f(x)$$

(or $F(x)$, $\phi(x)$, $\psi(x)$, ... as the case may be).

The nature of particular functions may, in very many cases, be illustrated and made easily intelligible as follows: draw two lines OX, OY at right angles to one another and produced indefinitely in both directions. We can represent values of x and y by distances measured from O along the lines OX, OY respectively, regard being paid, of course, to sign, and the positive directions of measurement being those indicated by arrows in Fig. 6.

Let a be any value of x for which y is defined and has (let us suppose) the single value b. Take $OA=a$, $OB=b$, and complete the rectangle $OAPB$. Imagine the point P marked on the diagram. This marking of the point P may be regarded as showing that the value of y for $x=a$ is b.

If to the value a of x corresponded *several* values of y (say b, b', b'') we should have, instead of the single point P, a number of points P, P', P''.

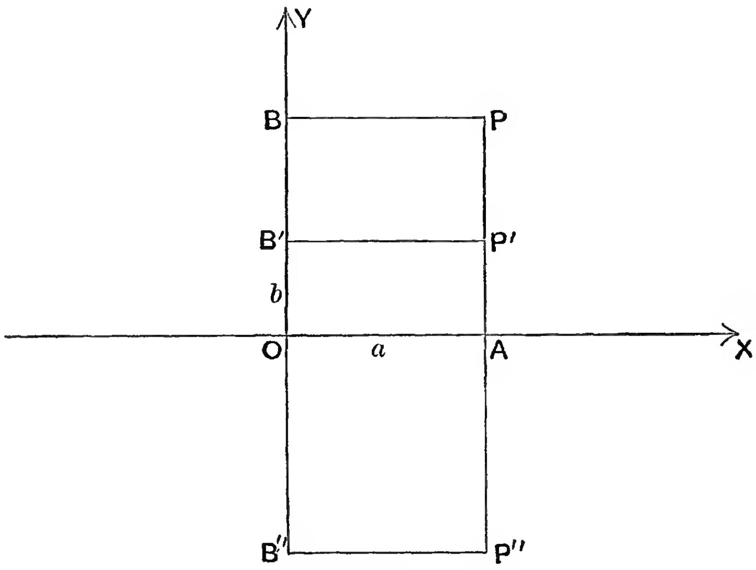


FIG. 6.

We shall call P the *point* (a, b); a and b the *coordinates of* P *referred to the axes* OX, OY; a the *abscissa*, b the *ordinate of* P; OX and OY the *axis of* x and the *axis of* y, or together the *axes of coordinates*, and O the *origin of coordinates*, or simply the *origin*.

Examples VIII. 1. Let P be the point (a, b), Q the point (α, β). Complete the parallelogram $OPRQ$. Show that R is the point $(a+\alpha, b+\beta)$.

2. The middle point of PQ is the point $\frac{1}{2}(a+\alpha)$, $\frac{1}{2}(b+\beta)$.

3. More generally, the line which divides PQ in the ratio $\mu : \lambda$ is the point $(\lambda a+\mu\alpha)/(\lambda+\mu)$, $(\lambda b+\mu\beta)/(\lambda+\mu)$. These expressions give, if the ratio $\mu : \lambda$ is properly chosen, the coordinates of any point on the line PQ.

4. The centre of mass of equal particles at the points (a_1, b_1), (a_2, b_2), ... (a_n, b_n) is the point $(a_1+a_2+\ldots+a_n)/n$, $(b_1+b_2+\ldots+b_n)/n$.

5. **Change of axes.** Draw through O lines OX', OY' making angles θ with OX, OY (Fig. 7). Draw PA', PB' perpendicular to OX', OY'. It is clear that P is determined if OA' and OB' are given just as much as if OA and OB are given. Let $OA=x$, $OB=y$, $OA'=x'$, $OB'=y'$. Then x' and y' are *the coordinates of* P *referred to the new axes* OX', OY'.

Prove that $x'=x\cos\theta+y\sin\theta$, $y'=-x\sin\theta+y\cos\theta$, and express x and y in terms of x' and y'.

6. In Ex. 5, the origin was left unchanged, but the new axes were inclined to the old ones. We might have taken new axes parallel to the old

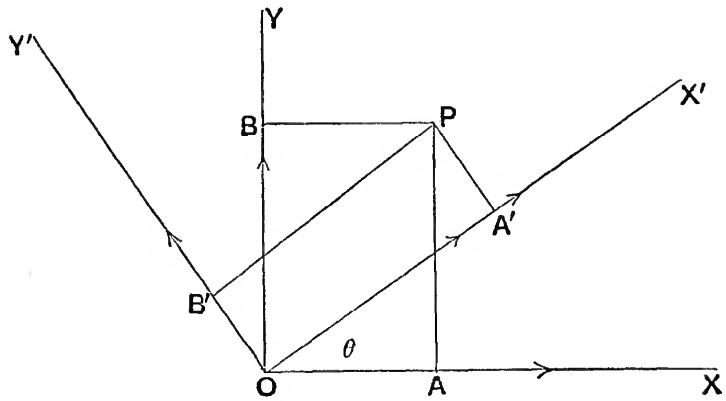


FIG. 7.

ones, but passing through a new origin O'. Let the coordinates of O' referred to the old axes be α, β. Express x' and y' in terms of x and y, and conversely.

7. A new origin O' is taken, and new axes $O'X'$, $O'Y'$ inclined at any angle to the old ones. Show, by means of the results of Exs. 5 and 6, that x' and y' may be expressed in terms of x and y by formulae of the type $x'=ax+by+c$, $y'=dx+ey+f$, where $a, b, \ldots$ are numbers independent of x and y.

11. The equation of a straight line. Let us now suppose that for all values a of x for which y is defined, the value b (or values $b, b', b'', \ldots$) of y, and the corresponding point P (or points $P, P', P'', \ldots$) have been determined. We call the aggregate of all these points the **graph** of the function y.

To take a very simple example, suppose that y is defined as a function of x by the equation

$$ax + by + c = 0 \quad \text{.......................(1)},$$

where a, b, c are any fixed numbers. Then y is a function of x which possesses all the characteristics (1), (2), (3) of § 9. It is easy to show that *the graph of y is a straight line.*

First suppose $a = 0$. Then y has the constant value $-c/b$, and the graph is obviously a straight line parallel to OX.

Next suppose a different from zero, and suppose that (x_1, y_1) (x_2, y_2) are any two points on the graph, so that

$$ax_1 + by_1 + c = 0, \quad ax_2 + by_2 + c = 0 \quad \text{...........(2)}.$$

The coordinates of any point P on the line joining (x_1, y_1) and (x_2, y_2) may (Ex. VIII. 3) be expressed in the form

$$\xi = (\lambda x_1 + \mu x_2)/(\lambda + \mu), \quad \eta = (\lambda y_1 + \mu y_2)/(\lambda + \mu).$$

But if we multiply the two equations (2) by λ and μ respectively, add the results, and divide by $\lambda + \mu$, we obtain

$$a\xi + b\eta + c = 0,$$

which shows that P lies on the graph. Hence the graph includes all the points of the line. And it cannot include any other points. For the line is not parallel to OX, since if it were y would be constant for all points on it, which is not the case. Hence there is one point on the line for which y has any value we like to assign. And so, if the graph contained a point (x', y') which did not lie on the line, there would be *two* values of x given by the equation $ax + by + c = 0$ when y had the value y': and this is obviously untrue. Thus the graph includes all the points of the line and no others.

We shall sometimes use another mode of expression. We shall say that when x and y vary in such a way that equation (1) is always true, *the locus of the point (x, y) is a straight line*, and we shall call (1) *the equation of the locus*, and say that the equation *represents* the locus. This use of the terms 'locus,' 'equation of the locus' is quite general, and may be applied whenever the graph of y is, in the ordinary sense of the word, a curve*, and the relation between x and y is capable of being represented by an analytical formula.

The preceding work does not apply when $b = 0$. The equation then reduces to $x = -c/a$, so that the distance of P from OY is constant—i.e. P lies on a line parallel to OY. In this case y does not occur in the equation at all, and so the latter cannot be regarded as defining y as a function of x. But it may be regarded as defining x as a function of y, viz. the constant $-c/a$.

The equation $ax + by + c = 0$ is *the general equation of the first degree*, for $ax + by + c$ is the most general polynomial in x and y which does not involve any terms of degree higher than the first

* 'Curve' of course includes straight line as a particular case. Some examples in which the 'graph' is *not*, in the ordinary sense of the word, a curve, will be found in Exs. XVI.

in x and y. Hence *the general equation of the first degree represents a straight line.* It is equally easy to prove the converse proposition, *the equation of any straight line is of the first degree.* After the discussion which precedes we may leave this as an exercise for the reader.

Examples IX. 1. The angles which the line $ax+by+c=0$ makes with OX are $\arctan(-a/b)$ and $\pi-\arctan(-a/b)$, where $\arctan\lambda$ denotes the numerically least angle whose tangent is λ.

2. If P is a point on the line, and x', y' are defined as in Ex. VIII. (5), show that

$$Ax'+By'+C=0,$$

where $$A=a\cos\theta+b\sin\theta,\quad B=b\cos\theta-a\sin\theta.$$

We call this equation *the equation of the line referred to the new axes OX', OY'*—it is the relation which connects the new coordinates x', y'. It will be observed that this equation also is of the first degree, as it obviously should be, since the proof that the equation of a straight line is of the first degree in no way depends upon what particular axes are chosen.

3. The coordinates of the point of intersection of $ax+by+c=0$ and $a'x+b'y+c'=0$ are

$$\frac{bc'-b'c}{ab'-a'b},\quad -\frac{ac'-a'c}{ab'-a'b};$$

unless $a/b=a'/b'$, in which case the lines are parallel.

4. The tangents of the angles between the lines in Ex. 3 are

$$\pm(ab'-a'b)/(aa'+bb'),$$

and the lines are perpendicular if $aa'+bb'=0$.

5. The length of the perpendicular from (ξ,η) on to $ax+by+c=0$ is

$$\frac{a\xi+b\eta+c}{\sqrt{(a^2+b^2)}},$$

the perpendicular being regarded as positive or negative according to the side of the line on which the point lies. [Positive when it is on the same side as O, if $c>0$: negative in the same circumstances if $c<0$.]

6. The equation $(ax+by+c)+\lambda(\alpha x+\beta y+\gamma)=0$ represents a line through the intersection of $ax+by+c=0$ and $\alpha x+\beta y+\gamma=0$, and, by proper choice of λ, may be made to represent any such line. Discuss the particular case in which $a/\alpha=\beta/b$.

7. Hence show how to find the equation of a line through the intersection of two given lines and parallel or perpendicular to a third.

8. The equation of the circle whose centre is (a, b) and radius r is

$$(x-a)^2+(y-b)^2=r^2.$$

Conversely, any equation of this form represents a circle.

9. The most general equation of the *second* degree in x and y, in which there is no term in xy, and x^2 and y^2 have equal coefficients, viz.

$$a(x^2+y^2)+2gx+2fy+c=0$$

represents a circle if $f^2+g^2>ac$. Discuss the cases in which $f^2+g^2 \leqq ac$.

10. Verify that the characteristic form of the equation of a circle (Ex. 9) is not altered by change of axes.

11. The general equation of a circle which passes through the points of intersection of two intersecting circles

$$(x-a)^2+(y-b)^2=r^2, \quad (x-a)^2+(y-\beta)^2=\rho^2$$

is

$$(x-a)^2+(y-b)^2-r^2+\lambda\,\{(x-a)^2+(y-\beta)^2-\rho^2\}=0.$$

12. If $\lambda=-1$ this last equation is of the first degree only, and represents the *common chord* of the two circles.

13. The two circles

$$x^2+y^2+2dx+2ey+k^2=0, \quad x^2+y^2+2\delta x+2\epsilon y+\kappa^2=0,$$

will represent a pair of intersecting circles if

$$d^2+e^2-k^2>0, \quad \delta^2+\epsilon^2-\kappa^2>0$$

and

$$4(d^2+e^2-k^2)(\delta^2+\epsilon^2-\kappa^2)>(2d\delta+2e\epsilon-k^2-\kappa^2)^2.$$

14. Show that the two circles in Ex. 13 will cut at right angles if

$$2d\delta+2e\epsilon=k^2+\kappa^2.$$

15. The area of the triangle formed by the points (x_1, y_1), (x_2, y_2), (x_3, y_3) is

$$\tfrac{1}{2}\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix},$$

taken positively. Hence deduce the result of § 11.

Examples X. 1. A point moves (a) so that its distance from a given line is constant, (b) so that its distances from two given lines are equal. Show that in each case the locus of the point is two straight lines (α) by geometrical reasoning, (β) by means of the results of § 11 and Ex. IX. 5.

2. The distances of a variable point P from a number of lines are p, p', p'', ..., and P moves so that

$$ap+bp'+cp''+\ldots=0$$

where a, b, c, ... are constants. Show that the locus of P consists of a number of straight lines.

3. A, B are fixed points, and P a variable point which moves so that (a) $\lambda \,.\, AP^2+\mu \,.\, BP^2=\text{const.}$, ($b$) $AP/BP=\text{const.}$ Show that the locus of P is in either case a circle.

4. A line of fixed length moves with its ends upon OX and OY. Find the equation of the locus of the point which divides the line in the ratio $\mu : \lambda$.

[Let the line be AB, meeting OX, OY in A, B, and let $OA=a$, $OB=b$. The coordinates of the point P in question (Ex. VIII. 3) are $\lambda a/(\lambda+\mu)$ and $\mu b/(\lambda+\mu)$. Also $a^2+b^2=\text{const.}=c^2$, say. Thus if $x=\lambda a/(\lambda+\mu)$, $y=\mu b/(\lambda+\mu)$, we deduce $\dfrac{x^2}{\lambda^2}+\dfrac{y^2}{\mu^2}=\dfrac{c^2}{(\lambda+\mu)^2}$.

If $\lambda=\mu$, i.e. if P is the middle point of AB, this is the equation of a circle.]

5. A line of constant length moves with its ends on a fixed circle. Prove that the locus of the point which divides the line in a fixed ratio is a concentric circle.

12. Polar coordinates. In what precedes we have determined the position of P by the lengths of $OM=x$, $MP=y$. If $OP=r$ and $MOP=\theta$, θ being an angle between 0 and 2π (measured in the positive direction), it is evident that

$$x=r\cos\theta, \qquad y=r\sin\theta,$$
$$r=\sqrt{(x^2+y^2)}, \qquad \cos\theta : \sin\theta : 1 :: x : y : r,$$

and that the position of P is equally well determined by a knowledge of r and θ. We call r and θ the *polar coordinates* of P. The former, it should be observed, is essentially positive.

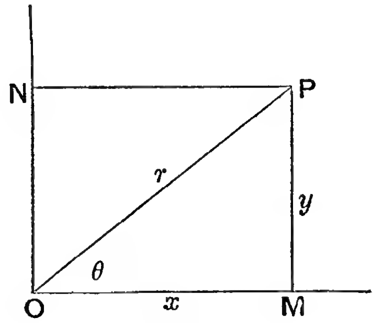


Fig. 8.

If P moves on a curve there will be some relation between r and θ, say $r=f(\theta)$ or $\theta=F(r)$. This we call the *polar equation* of the locus. The polar equation may be deduced from the (x, y) equation (or *vice versa*) by the formulae above.

It should be observed that (x, y) and (r, θ) are only two out of an infinite variety of 'systems of coordinates' which may be used to fix the position of P.

Examples XI. 1. The polar equation of a straight line is of the form

$$r\cos(\theta-\alpha)=p,$$

where p and α are constants.

2. The equation $r=2a\cos\theta$ represents a circle passing through the origin. So do $r=2a\sin\theta$ or $r=\lambda\cos\theta+\mu\sin\theta$. Find the radius of each circle.

3. The general equation of a circle is of the form

$$r^2+c^2-2rc\cos(\theta-\alpha)=a^2,$$

where a, c, and α are constants.

13. Further examples of functions and their graphical representation. In all of Exs. IX. and X. we were concerned with two very simple functions of x, viz. the functions y defined by the equations $\alpha x + \beta y + \gamma = 0$ or $(x-a)^2 + (y-b)^2 = r^2$. Only in Ex. X. 4 did we meet for a moment a slightly more general type of functional relation. The examples which follow will give the reader a better notion of the infinite variety of possible types of functions.

A. **Polynomials.** The meaning of the term *polynomial in x* was explained in Ch. I. It denotes a function of the form

$$a_0 x^m + a_1 x^{m-1} + \dots + a_m$$

where $a_0, a_1, \dots a_m$ are constants. The simplest polynomials are the *simple powers*

$$y = x,\ x^2,\ x^3,\ \dots,\ x^m,\ \dots.$$

The graph of the function x^m is of two distinct types, according as m is even or odd.

First let $m = 2$. Then three points on the graph are

$$(0, 0),\quad (1, 1),\quad (-1, 1).$$

Any number of additional points on the graph may be found by assigning other special values to x: thus the values

$$x = \tfrac{1}{2},\ 2,\ 3,\ -\tfrac{1}{2},\ -2,\ 3$$

give $$y = \tfrac{1}{4},\ 4,\ 9,\quad \tfrac{1}{4},\quad 4,\ 9.$$

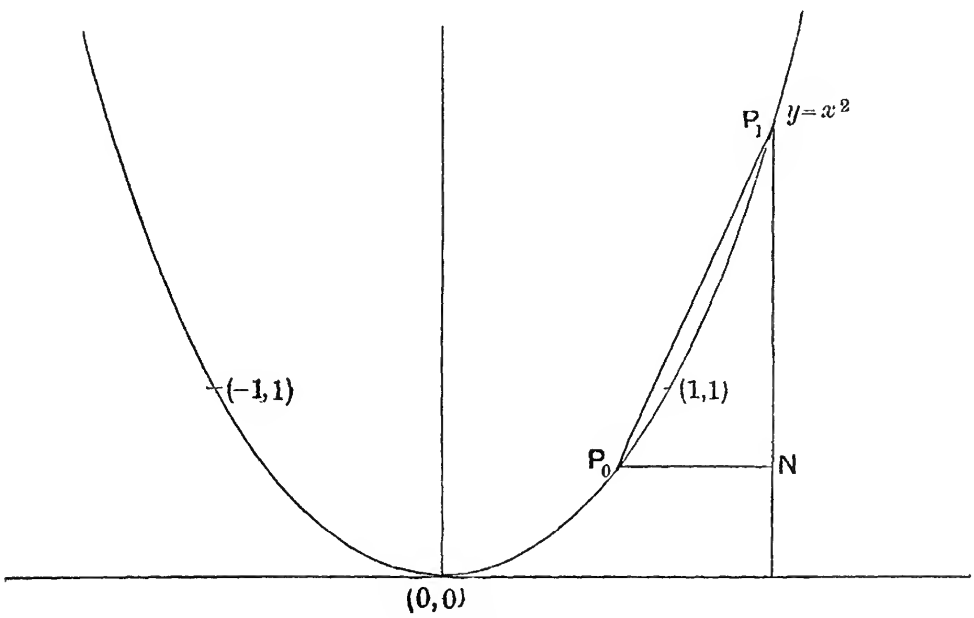


Fig. 9.

If the reader will plot off a fair number of points on the graph he will be led to conjecture that the form of the graph is something like that shown in Fig. 9. If he draws a curve through the special points which he has proved to lie on the graph and then tests its accuracy by giving x new values, and calculating the corresponding values of y, he will find that they lie as near to the curve as it is reasonable to expect, when the inevitable inaccuracies of drawing are considered.

There is, however, one fundamental question which we cannot answer adequately at present. The reader has no doubt some notions as to what is meant by a *continuous* curve, a curve without breaks or jumps—such a curve, in fact, as is roughly represented in Fig. 9. The question is whether the graph of the function $y=x^2$ is in fact such a curve. This cannot be *proved* by merely constructing any number of isolated points on the curve, although the more such points we construct the more probable it will appear.

This question cannot be discussed properly until Ch. IV. In that chapter we shall consider in detail what our common sense idea of continuity really means, and how we can prove that such graphs as the one now considered, and others which we shall consider later on in this chapter, are really continuous curves. For the present the reader may be content to draw his curves as common sense dictates.

It is easy to see that the curve $y=x^2$ is everywhere *convex* to the axis of x. Let P_0, P_1 (Fig. 9) be the points (x_0, x_0^2), (x_1, x_1^2). Then

$$\tan NP_0P_1=\frac{NP_1}{P_0N}=\frac{x_1^2-x_0^2}{x_1-x_0}=x_0+x_1,$$

and, if P_0 is kept fixed, this increases as x_1 increases—i.e. the slope of P_0P_1 becomes steeper and steeper.

The curve $y=x^4$ is similar to $y=x^2$ in general appearance, but flatter near O, and steeper beyond the points A, A' (Fig. 10). And $y=x^m$, where m is even and greater than 4, is still more so. And as m gets larger and larger the flatness and steepness grow more and more pronounced, until the curve is practically indistinguishable from the thick broken line in the figure.

The reader should next consider the curves given by $y=x^m$, when m is odd. The fundamental difference between the two

cases is that whereas when m is even $(-x)^m = x^m$, so that the curve is symmetrical about OY, when m is odd $(-x)^m = -x^m$, so that y is negative when x is negative. Fig. 11 shows the curves

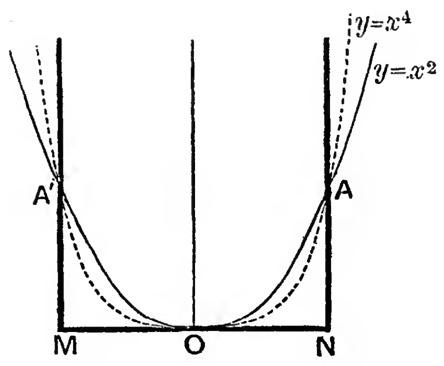


FIG. 10.

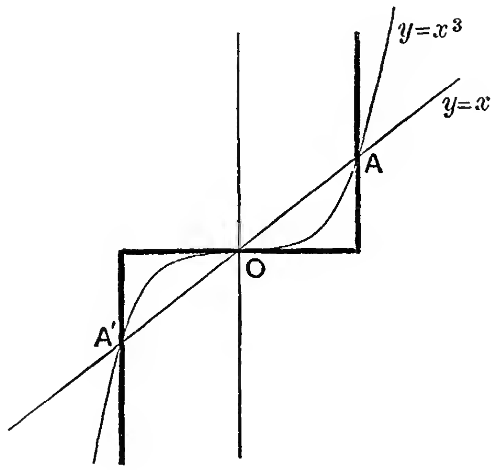


FIG. 11.

$y = x$, $y = x^3$, and the form to which $y = x^m$ approximates for larger odd values of m.

It is now easy to see how (theoretically at any rate) the graph of any polynomial may be constructed. In the first place, from the graph of $y = x^m$ we can at once derive that of Cx^m, where C is a constant, by multiplying the ordinate of every point of the curve by C. And if we know the graphs of $f(x)$ and $F(x)$ we can find that of $f(x) + F(x)$ by taking the ordinate of every point to be the sum of the ordinates of the corresponding points on the two original curves.

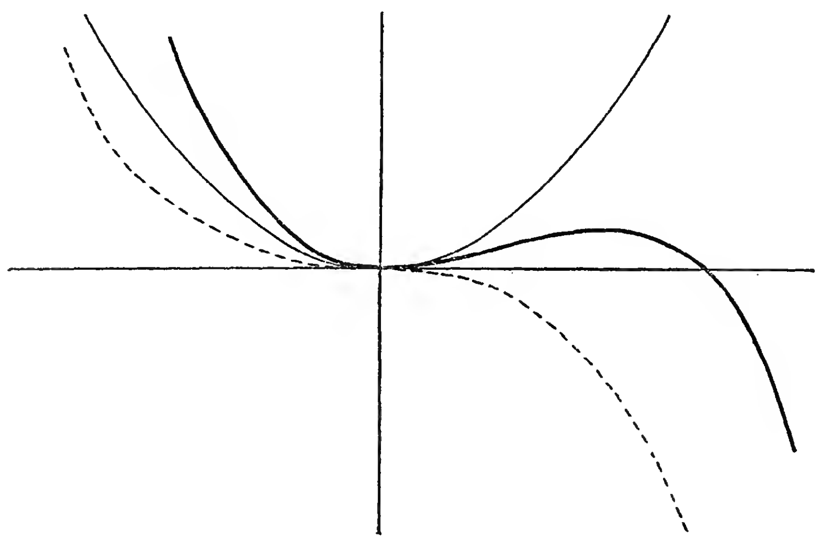


FIG. 12.

Fig. 12 shows the graph of $y=2x^2-x^3$, constructed in this way. The thin line is $y=2x^2$; the dotted line $y=-x^3$. In order to prevent the figure becoming of an awkward size, the scale for measurements along the axis of y has been taken to be one-quarter of that for measurements along the axis of x. This is often convenient: of course any ratio of the scales may be chosen.

The drawing of graphs of polynomials is however so much facilitated by the use of more advanced methods, which will be explained later on, that we shall not pursue the subject further here.

Examples XII. 1. Trace the curves $y=7x^4$, $y=3x^5$, $y=x^{10}$.

[The reader should draw the curves carefully, choosing the scales of measurement along OX and OY so as to get a convenient figure: but all three curves should be drawn *in one figure*. The reader will then realise how rapidly the higher powers of x increase, as x gets larger and larger, and will see that, in such a polynomial as

$$x^{10}+3x^5+7x^4,$$

(or even $x^{10}+30x^5+700x^4$) it is the *first* term which is of really preponderant importance when x is fairly large. Thus even when $x=4$, $x^{10}>1{,}000{,}000$, while $30x^5<35{,}000$ and $700x^4<180{,}000$; while if $x=10$ the preponderance of the first term is still more marked.]

2. Compare the relative magnitudes of x^{12}, $1{,}000{,}000x^6$, $1{,}000{,}000{,}000{,}000x$ when $x=1$, 10, 100, etc.

[The reader should make up a number of examples of this type for himself. This idea of the *relative rate of growth* of different functions of x is one with which we shall often be concerned in the following chapters.]

3. Draw the graph of $ax^2+2bx+c$.

[Here $y-\{(ac-b^2)/a\}=a\{x+(b/a)\}^2$. If we take new axes parallel to the old and passing through the point $-b/a$, $(ac-b^2)/a$, the new equation is $y'=ax'^2$. The reader should consider a few different cases in which a, b, c have numerical values, sometimes positive and sometimes negative.]

4. Trace the curves $y=x^3-3x+1$, $y=x^2(x-1)$, $y=x(x-1)^2$.

14. B. Rational Functions. The class of functions which ranks next to that of polynomials in simplicity and importance is that of rational functions. In Ch. I. we defined a rational function as the quotient of one polynomial by another: thus if $P(x)$, $Q(x)$ are polynomials we may denote the general rational function by

$$R(x)=\frac{P(x)}{Q(x)}.$$

In the particular case when $Q(x)$ reduces to unity or any other constant (i.e. does not involve x), $R(x)$ reduces to a polynomial: thus the class of rational functions includes that of polynomials as a sub-class. The following points concerning the definition should be noticed.

(1) We usually suppose that $P(x)$ and $Q(x)$ have no common factor $x+a$ or $x^p+ax^{p-1}+bx^{p-2}+\dots+k$, all such factors being removed by division.

(2) It should however be observed that this removal of common factors *does as a rule change the function*. Consider for example the function x/x, which is a rational function. On removing the common factor x we obtain $1/1=1$. But the original function is not *always* equal to 1: it is equal to 1 only so long as $x \neq 0$. If $x=0$ it takes the form $0/0$, which is meaningless. Thus the function x/x is equal to 1 if $x \neq 0$ and is undefined when $x=0$. It therefore differs from the function 1 which is *always* equal to 1.

(3) Such a function as

$$\left(\frac{1}{x+1}+\frac{1}{x-1}\right)\Big/\left(\frac{1}{x}+\frac{1}{x-2}\right)$$

may be reduced, by the ordinary rules of algebra, to the form

$$\frac{x^2(x-2)}{(x-1)^2(x+1)},$$

which is a rational function of the standard form. But here again it must be noticed that the reduction is not *always* legitimate. In order to calculate the value of a function for a given value of x we must substitute the value for x in the function *in the form in which it is given*. In the case of this function the values $x=-1$, 1, 0, 2 all lead to a meaningless expression, and so the function is not defined for these values. The same is true of the reduced form, so far as the values ± 1 are concerned. But $x=0$ or 2 gives the value 0. Thus once more the two functions are not strictly equivalent.

(4) But, as appears from the particular example considered under (3), even when the function has been reduced to a rational function of the standard form there will generally be a certain number of values of x for which it is not defined. These are the values of x (if any) for which the denominator vanishes. Thus $(x^2-7)/(x^2-3x+2)$ is not defined when $x=1$ or 2.

(5) Generally we agree, in dealing with expressions such as those considered in (2) and (3), to disregard the exceptional values of x for which such processes of simplification as were used there are illegitimate, and to reduce our function to the standard form of rational function. The reader will easily verify that (on this understanding) the sum, product, or quotient of two rational functions may themselves be reduced to rational functions of the standard type. And generally *a rational function of a rational function is itself a rational function*: i.e. if in $z=P(y)/Q(y)$, where P and Q are

polynomials, we substitute $y=P_1(x)/Q_1(x)$, we obtain on simplification an equation of the form $z=P_2(x)/Q_2(x)$.

(6) It is in no way presupposed in the definition of a rational *function* that the *constants* which occur as coefficients should be rational *numbers*. The word rational has reference solely to the way in which the variable x appears in the function. Thus

$$\frac{x^2+x+\sqrt{3}}{x\sqrt[3]{2}-\pi}$$

is a rational function.

The use of the word rational arises as follows. The rational function $P(x)/Q(x)$ may be generated from x by a definite number of operations upon x, including only multiplication of x by itself or a constant, addition of terms thus obtained, and division of one function, obtained by such multiplications and additions, by another. In so far as the variable x is concerned, this procedure is very much like that by which all rational numbers can be obtained from unity, a procedure exemplified in the equation

$$\frac{5}{3}=\frac{1+1+1+1+1}{1+1+1}.$$

Again, *any* function which can be deduced from x by the elementary operations mentioned above, using at each stage of the process functions which have already been obtained from x in the same way, can be reduced to the standard type of rational function. The most general kind of function which can be obtained in this way is sufficiently illustrated by the example

$$\left(\frac{x}{x^2+1}+\frac{2x+7}{x^2+\dfrac{11x-3\sqrt{2}}{9x+1}}\right)\Bigg/\left(17+\frac{2}{x^3}\right),$$

which can obviously be reduced to the standard type of rational function.

15. The drawing of graphs of rational functions, even more than that of polynomials, is immensely facilitated by the use of

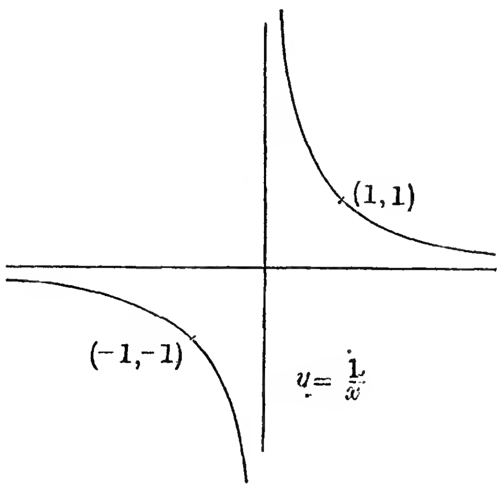


FIG. 13.

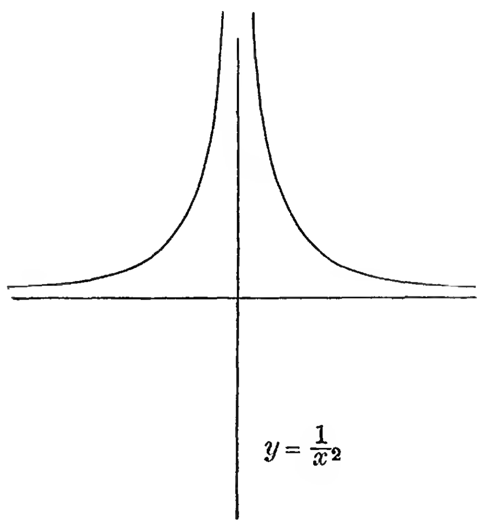


FIG. 14.

methods depending upon the differential calculus. We shall therefore content ourselves at present with a very few examples.

Examples XIII. 1. Draw the graphs of $y=1/x$, $y=1/x^2$, $y=1/x^3$, ...

[The figure shows the graphs of the first two curves. It should be observed that, since $1/0$, $1/0^2$, ... are meaningless expressions, these functions are not defined for $x=0$.]

2. Trace $y=x+(1/x)$, $x-(1/x)$, $x^2+(1/x^2)$, $x^2-(1/x^2)$ and $ax+(b/x)$, taking various values, positive and negative, for a and b.

3. Trace
$$y=\frac{x+1}{x-1},\quad \left(\frac{x+1}{x-1}\right)^2,\quad \frac{x^2+1}{x^2-1},\quad \frac{1}{(x-1)^2},\quad (x+1)^2-\frac{1}{(x-1)^2}.$$

4. Trace $y=1/(x-a)(x-b)$, $1/(x-a)(x-b)(x-c)$, where $a<0<b<c$.

5. Sketch the general form assumed by the curves $y=1/x^m$ as m becomes larger and larger, considering separately the cases in which m is odd or even.

16. C. Explicit Algebraical Functions. The next important class of functions is that of *explicit algebraical functions.* These are functions which can be generated from x by a definite number of operations such as those used in generating rational functions, together with a definite number of operations of root extraction. Thus

$$\frac{\sqrt{(1+x)}-\sqrt[3]{(1-x)}}{\sqrt{(1+x)}+\sqrt[3]{(1-x)}},\qquad \sqrt{x}+\sqrt{(x+\sqrt{x})},$$
$$\left(\frac{x^2+x+\sqrt{3}}{x\sqrt[3]{2}-\pi}\right)^{\frac{2}{3}},$$

are explicit algebraical functions, and so is $x^{m/n}$ (i.e. $\sqrt[n]{x^m}$) where m and n are any integers.

Functions such as these differ fundamentally from rational functions in two respects. In the first place, a rational function is always defined for all values of x with a certain number of isolated exceptions. But such a function as $\sqrt{x}$ is undefined for *a whole range* of values of x (i.e. all negative values). Secondly, the function, when x has a value for which it is defined, has generally several values. Thus, if $x>0$, $\sqrt{x}$ has *two* values, of opposite signs.

Examples XIV. 1. $\sqrt{\{(x-a)(b-x)\}}$, where $a<b$, is defined only for $a \leqq x \leqq b$. If $a<x<b$ it has two values: if $x=a$ or b only one, viz. 0.

2. Consider similarly $\sqrt{(x-a)(x-b)(c-x)}$ $(a<b<c)$,

$$\sqrt{x(x^2-a^2)},\quad \sqrt[3]{(x-a)^2(b-x)}\quad (a<b),$$

$$\frac{\sqrt{1+x}-\sqrt{1-x}}{\sqrt{1+x}+\sqrt{1-x}},\quad \sqrt{\{x+\sqrt{x}\}}.$$

3. Trace $y=\sqrt{x}$, $\sqrt[3]{x}$, $\sqrt[3]{x^2}$, $(1+\sqrt{x})/(1-\sqrt{x})$.

4. Trace $y=\sqrt{(a^2-x^2)}$, $y=b\sqrt{\{1-(x^2/a^2)\}}$.

17. D. **Implicit Algebraical Functions.** It is easy to verify that if

$$y=\frac{\sqrt{(1+x)}-\sqrt[3]{(1-x)}}{\sqrt{(1+x)}+\sqrt[3]{(1-x)}}$$

then

$$\left(\frac{1+y}{1-y}\right)^6=\frac{(1+x)^3}{(1-x)^2}$$

or if

$$y=\sqrt{x}+\sqrt{(x+\sqrt{x})}$$

then

$$y^4-(4y^2+4y+1)x=0.$$

Each of these equations is of the form

$$y^m+R_1y^{m-1}+\dots+R_m=0 \quad\text{.................(1)},$$

where R_1, R_2, ..., R_m are rational functions of x: and the reader will easily verify that, if y is any one of the functions considered in the last set of examples, y satisfies an equation of this form. It is naturally suggested that the same is true of any explicit algebraic function. And this is in fact true, and indeed not difficult to prove, though we shall not delay to write out a formal proof here.

An example should make clear to the reader the lines on which such a proof would proceed. Let

$$y=\frac{x+\sqrt{x}+\sqrt{\{x+\sqrt{x}\}}+\sqrt[3]{(1+x)}}{x-\sqrt{x}+\sqrt{\{x+\sqrt{x}\}}-\sqrt[3]{(1+x)}}.$$

Then we have the equations

$$y=\frac{x+u+v+w}{x-u+v-w},$$

$$u^2=x,\quad v^2=x+u,\quad w^3=1+x,$$

and we have only to eliminate u, v, w between these equations in order to obtain an equation of the form desired.

We are therefore led to give the following definition: *a function $y=f(x)$ will be said to be an algebraical function of x if it is the root of an equation such as* (1), *i.e. the root of an equation of the m^{th} degree in y, whose coefficients are rational functions of x.*

This class of functions includes all the explicit algebraical functions considered in § 16. But it also includes other functions which cannot be expressed as explicit algebraical functions. For it is known that such an equation as (1) cannot *as a rule* be solved explicitly for y in terms of x, when m is greater than 4, though such a solution is always possible if $m=1$, 2, 3, or 4 and in special cases for higher values of m.

The definition of an algebraical function should be compared with that of an algebraical number given in the last chapter (Misc. Exs. 30).

Examples XV. 1. If $m=1$, y is a *rational* function.

2. If $m=2$ the equation is $y^2+R_1y+R_2=0$, so that

$$y=\tfrac{1}{2}\{-R_1\pm\surd(R_1^2-4R_2)\}.$$

This function is defined for all values of x for which $R_1^2\geqq 4R_2$. It has two values if $R_1^2>4R_2$ and one if $R_1^2=4R_2$.

If $m=3$ or 4 we can use the methods explained in treatises on Algebra for the solution of cubic and biquadratic equations. But as a rule the process is complicated and the results inconvenient in form, and we can generally study the properties of the function better by means of the original equation.

3. Consider the functions defined by the equations

$$y^2-2y-x^2=0,\quad y^2-2y+x^2=0,\quad y^4-2y^2+x^2=0,$$

in each case obtaining y as an explicit function of x, and stating for what values of x it is defined.

4. Find algebraical equations, with coefficients rational in x, satisfied by each of the functions

$$\surd x+\surd(1/x),\quad \sqrt[3]{x}+\sqrt[3]{(1/x)},\quad \sqrt[4]{x}+\sqrt[4]{(1/x)},\quad \sqrt[3]{(1+x)}+\sqrt[3]{(1-x)},$$
$$\surd\{x+\surd x\},\quad \surd[x+\surd\{x+\surd x\}].$$

5. Consider the equation $\qquad y^4=x^2.$

[Here $y^2=\pm x$. If x is positive $y=\pm\surd x$: if negative $y=\pm\surd(-x)$. Thus the function has two values for all values of x save $x=0$, when it has the one value 0.]

6. An algebraical function of an algebraical function of x is itself an algebraical function of x.

[For we have

$$y^m + R_1(z)\, y^{m-1} + \dots + R_m(z) = 0,$$

where

$$z^n + S_1(x)\, z^{n-1} + \dots + S_n(x) = 0.$$

Eliminating z we find an equation of the form

$$y^p + T_1(x)\, y^{p-1} + \dots + T_p(x) = 0.$$

Here all the capital letters denote rational functions.]

7. An example should perhaps be given of an algebraical function which *cannot* be expressed in an explicit algebraical form. Such an example is the function y defined by the equation

$$y^5 - y - x = 0.$$

But a proof that we cannot find an explicit algebraical expression for y in terms of x is difficult, and cannot be attempted here.

18. Transcendental Functions. All functions of x which are not rational or even algebraical are called *transcendental* functions. This class of functions, being defined in so purely negative a manner, naturally includes an infinite variety of whole kinds of functions of varying degrees of simplicity and importance. Among these we can at present distinguish two kinds which are particularly interesting.

E. **The direct and inverse trigonometrical or circular functions.** These are the sine and cosine functions of elementary trigonometry, and their inverses, and the functions derived from them. We may assume that the reader is familiar with their most important properties.

Examples XVI. 1. Draw the graphs of $\sin x$, $\cos x$, and $a\cos x + b\sin x$.

[Since $a\cos x + b\sin x = \beta\cos(x-\alpha)$, where $\beta = \sqrt{(a^2+b^2)}$, and α is an angle whose cosine and sine are $a/\sqrt{(a^2+b^2)}$ and $b/\sqrt{(a^2+b^2)}$, the graphs of these three functions are similar in character.]

2. Draw the graphs of $\cos^2 x$, $\sin^2 x$, $a\cos^2 x + b\sin^2 x$.

3. Suppose the graphs of $f(x)$ and $F(x)$ drawn. Then the graph of

$$f(x)\cos^2 x + F(x)\sin^2 x$$

is a wavy curve which oscillates between the curves $y = f(x)$, $y = F(x)$. Draw the graph when $f(x)$, $F(x)$ are any pair of the functions

$$1/x^2, \quad 1/x, \quad 1, \quad x, \quad x^2, \quad ax+b, \quad x+(1/x).$$

4. Discuss in the same manner the form of the graph of

$$f(x)\cos x + F(x)\sin x.$$

5. Draw the graphs of $x+\sin x$, $x^2+\sin x$, $(1/x)+\sin x$, $x\sin x$, $x^2\sin x$, $(\sin x)/x$.

6. Draw the graph of $\sin(1/x)$.

[If $y=\sin(1/x)$, $y=0$ when $x=1/m\pi$, where m is any integer. Similarly $y=1$ when $x=1/(2m+\frac{1}{2})\pi$ and $y=-1$ when $x=1/(2m-\frac{1}{2})\pi$. The curve is entirely comprised between the lines $y=\pm 1$. It oscillates up and down, the rapidity of the oscillations becoming greater and greater as x approaches 0. For $x=0$ the function is undefined. When x is large y is small. The negative half of the curve is an inversion of the positive half (Fig. 15).]

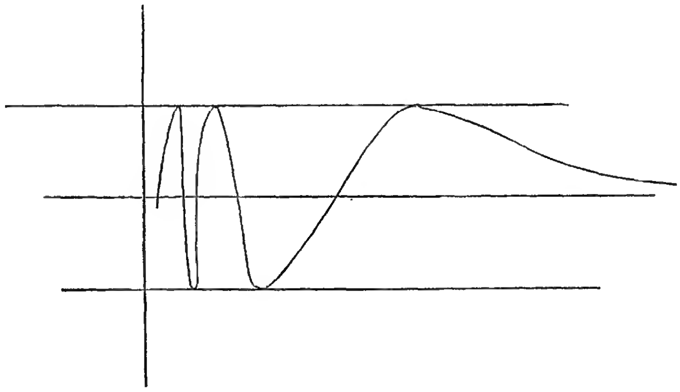


FIG. 15.

7. Draw the graph of $x\sin(1/x)$.

[This curve is comprised between the lines $y=\pm x$ just as the last curve was comprised between the lines $y=\pm 1$ (Fig. 16).]

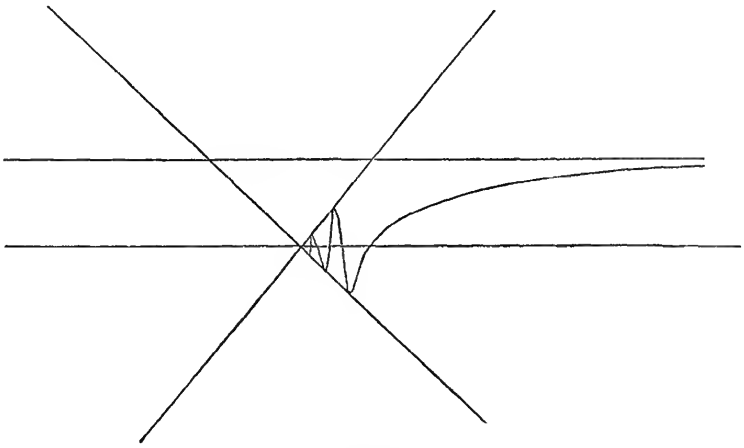


FIG. 16.

8. Draw the graphs of $x^2\sin(1/x)$, $(1/x)\sin(1/x)$, $\sin^2(1/x)$, $\{x\sin(1/x)\}^2$, $a\cos^2(1/x)+b\sin^2(1/x)$, $\sin x+\sin(1/x)$, $\sin x\sin(1/x)$.

9. Draw the graphs of $\cos x^2$, $\sin x^2$, $a\cos x^2+b\sin x^2$.

10. Draw the graphs of $\operatorname{arc}\cos x$ and $\operatorname{arc}\sin x$.

[If $y = \text{arc}\cos x$, $x = \cos y$. This enables us to draw the graph of x, considered as a function of y, and the same curve shows y as a function of x. It is clear that y is only defined for $-1 \leqq x \leqq 1$, and is infinitely many valued for these values of x. As the reader no doubt remembers, there is, when $-1 < x < 1$, a value of y between 0 and π, say α, and the other values of y are given by the formula $2n\pi \pm \alpha$, where n is any integer, positive or negative.]

11. Draw the graphs of

$$\tan x, \quad \cot x, \quad \sec x, \quad \text{cosec}\, x, \quad \tan^2 x, \quad \cot^2 x, \quad \sec^2 x, \quad \text{cosec}^2 x.$$

12. Draw the graphs of $\text{arc}\tan x$, $\text{arc}\cot x$, $\text{arc}\sec x$, $\text{arc}\,\text{cosec}\, x$. Give formulae expressing all the values of each of these functions in terms of any particular value.

13. Draw the graphs of $\tan(1/x)$, $\cot(1/x)$, $\sec(1/x)$, $\text{cosec}(1/x)$.

14. Show that $\sin x$ and $\cos x$ are not rational functions of x.

[It is easy to see that no function which, like the sine or cosine, has a *period*, can possibly be a rational function. For suppose that

$$f(x) = P(x)/Q(x),$$

where P and Q are polynomials, and $f(x) = f(x + 2\pi)$, each of these equations holding for all values of x. Let $f(0) = k$. Then the equation

$$P(x) - kQ(x) = 0$$

is satisfied by an infinite number of values of x, viz. $x = 0$, 2π, 4π, etc., and so it is an *identity*. Thus $f(x) = k$ for all values of x, i.e. $f(x)$ is a mere constant.]

15. Show, more generally, that no function with a period can be an algebraical function of x.

[Let the equation which defines the algebraical function be

$$y^m + R_1 y^{m-1} + \ldots + R_m = 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1),$$

where $R_1, \ldots$ are rational functions of x. This may be put in the form

$$P_0 y^m + P_1 y^{m-1} + \ldots + P_m = 0,$$

where $P_0, P_1, \ldots$ are polynomials in x. Arguing as above we see that

$$P_0 k^m + P_1 k^{m-1} + \ldots + P_m = 0$$

is an identity. Hence $y = k$ satisfies the equation (1) for all values of x, and one set of values of our algebraical function reduces to a constant.

Now divide (1) by $y - k$ and repeat the argument m times. Our final conclusion is that our algebraical function has, for any value of x, the same m values $k, k', \ldots$; i.e. it is composed of m mere constants.]

16. The inverse sine and inverse cosine are not rational or algebraical functions.

[This follows from the fact that for any value of x between -1 and $+1$, $\text{arc}\sin x$ and $\text{arc}\cos x$ have infinitely many values.]

19. F. Other classes of transcendental functions. Next in importance to the trigonometrical functions come the exponential and logarithmic functions, which will be discussed in Chh. IX. and X. But these functions are beyond our range at present. And most of the other classes of transcendental functions whose properties have been studied, such as the elliptic functions, Bessel's and Legendre's functions, Gamma-functions, and so forth, lie altogether beyond the range of this book. There are however some elementary types of functions which, though of much less importance theoretically than the rational, algebraical, or trigonometrical functions, are particularly instructive as illustrations of the possible varieties of the functional relation.

Examples XVII. 1. Let $y=[x]$, where $[x]$ denotes the algebraically greatest integer contained in x. The graph is shown in Fig. 17 (a).

2. $y=x-[x]$. (Fig. 17 (b).)

3. $y=\sqrt{\{x-[x]\}}$. (Fig. 17 (c).)

4. $y=[x]+\sqrt{\{x-[x]\}}$. (Fig. 17 (d).)

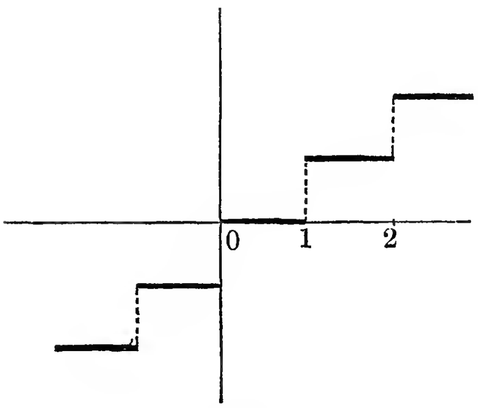


FIG. 17 a.

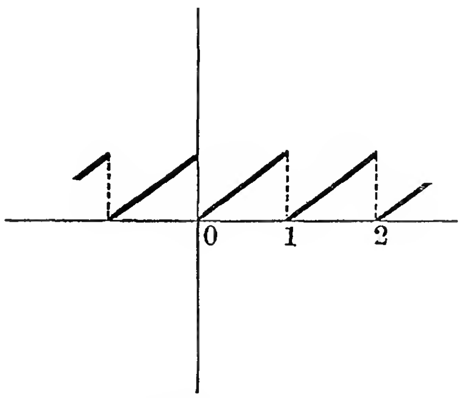


FIG. 17 b.

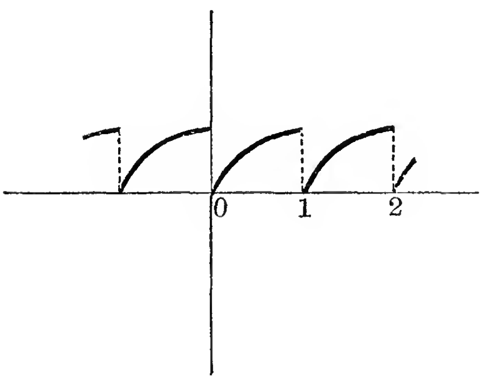


FIG. 17 c.

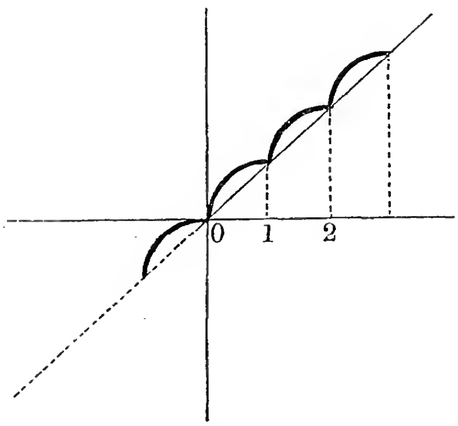


FIG. 17 d.

5. $y=(x-[x])^2$, $[x]+(x-[x])^2$.

6. $y=[\sqrt{x}]$, $[x^2]$, $\sqrt{x}-[\sqrt{x}]$, $x^2-[x^2]$, $[1-x^2]$.

7. Let y be defined as *the largest prime factor of* x (cf. Exs. VII. 6). Then y is defined only for integral values of x. When

$$\begin{array}{llllllllllllll} x= & \pm 1, & 2, & 3, & 4, & 5, & 6, & 7, & 8, & 9, & 10, & 11, & 12, & 13, \ldots \\ y= & 1, & 2, & 3, & 2, & 5, & 3, & 7, & 2, & 3, & 5, & 11, & 3, & 13, \ldots \end{array}$$

The graph consists of a number of isolated points.

8. Let y be *the denominator of* x (Exs. VII. 7). In this case y is defined only for rational values of x. We can mark off as many points on the graph as we please, but the result is not in any ordinary sense of the word a curve, and there are no points corresponding to any irrational values of x.

Draw the straight line joining the points $(N-1, N)$, (N, N). Show that the number of points of the locus which lie on this line is equal to the number of numbers less than and prime to N.

9. Let $y=0$ when x is an integer, $y=x$ when x is not an integer. The graph is derived from the straight line $y=x$ by taking out the points

$$\ldots(-1,-1),\quad (0,0),\quad (1,1),\quad (2,2),\ \ldots,$$

and adding the points $(-1, 0)$, $(0, 0)$, $(1, 0)$, ... on the axis of x.

The reader may possibly regard this as an *unreasonable* function. *Why*, he may ask, if y is equal to x for all values of x save integral values, should it not be equal to x for integral values too? The answer is simply, *why should it?* The function y does in point of fact answer to the definition of a function: there is a relation between x and y such that when x is known y is known. We are perfectly at liberty to take this relation to be what we please, however arbitrary and apparently futile. This function y is, of course, a quite different function from that one which is *always* equal to x, whatever value, integral or otherwise, x may have. Let us take an apparently still more arbitrary example.

10. Let
$$\begin{aligned} y&=0 \text{ when } x=-2\tfrac{1}{2}, \\ y^2&=1 \text{ when } x=-1, \\ y&=\sin x \text{ if } -\tfrac{1}{2}\leqq x\leqq \tfrac{1}{2}, \end{aligned}$$
and
$$y^2=x^2 \quad \text{if} \quad 1<x\leqq 2,$$
except that $y=-1$ when $x=1\frac{1}{2}$. And for $x=3$ let y have all values between -1 and $+1$. Finally, suppose that y is not defined at all except for the various values just enumerated. The graph is shown in Fig. 18. It consists of the curved arc L, the line P, the lines M and N, from which however the middle points and the ends nearest the axis of x must be taken out, and the four isolated points A, B, C, D. We notice further that y has *infinitely many* values for $x=3$, *two* for $x=-1$ and $1<x<1\frac{1}{2}$ and $1\frac{1}{2}<x\leqq 2$, *one* for $x=-2\frac{1}{2}$, $-\frac{1}{2}\leqq x\leqq\frac{1}{2}$ and $x=1\frac{1}{2}$, and *none* for any other value of x.

This example is given merely to illustrate *possibilities*; it is not suggested that such functions as these are likely to be of any practical importance.

The reader should however not be too ready to assume that even from the practical point of view it is only what is obvious and straightforward which is

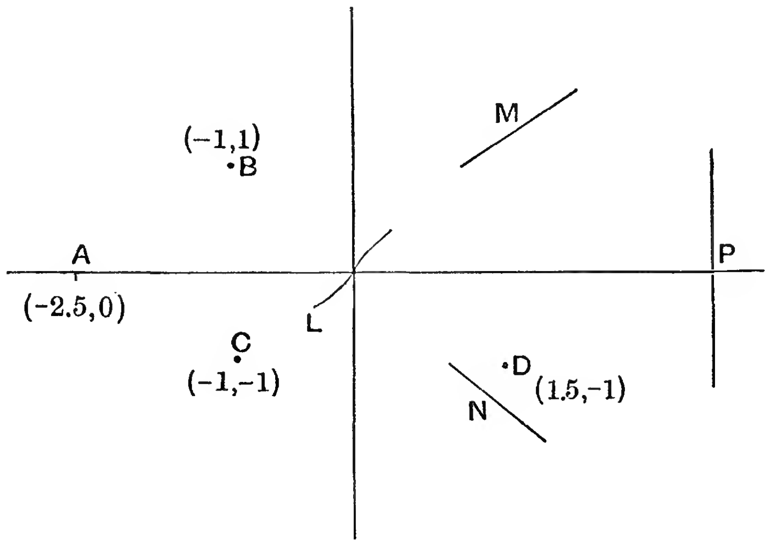


Fig. 18.

important. If he turns back to Exs. VII. 4, 5, for instance, he will see examples of functions suggested by physical considerations and defined by different formulae for different ranges of values of x.

11. Let $y=1$ when x is rational, but $y=0$ when x is irrational. The graph consists of two series of points arranged upon the lines $y=1$ and $y=0$. To the eye it is not distinguishable from two continuous straight lines, but in reality an infinite number of points are missing from each line.

12. Let $y=x$ when x is irrational and $y=\sqrt{\{(1+p^2)/(1+q^2)\}}$ when x is a rational fraction p/q.

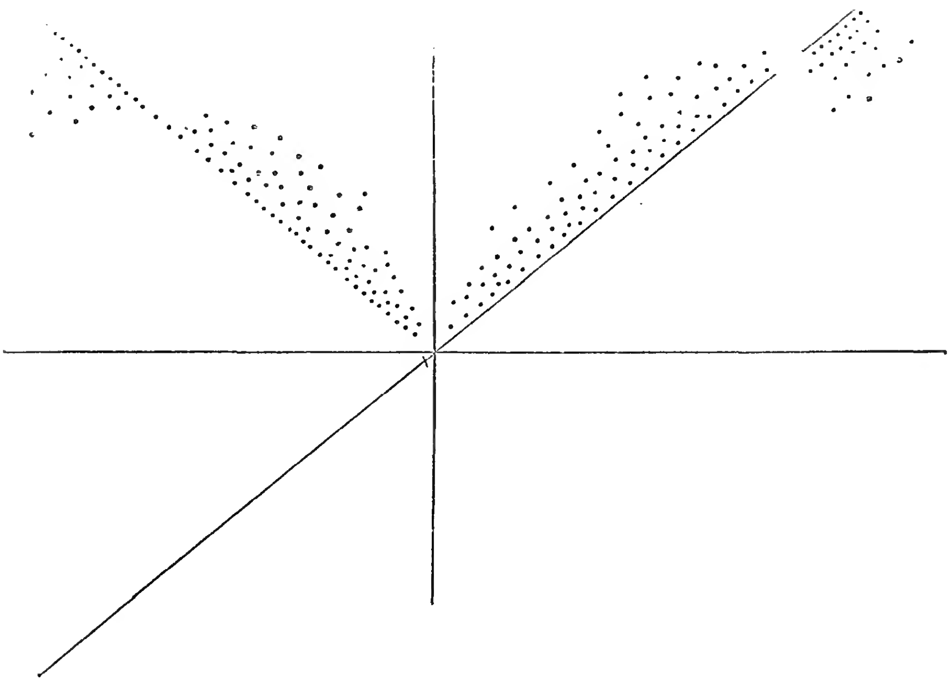


Fig. 19.

The irrational values of x contribute to the graph a curve in reality discontinuous, but apparently not to be distinguished from the straight line $y=x$.

Now consider the rational values of x. First let x be positive. Then $\sqrt{\{(1+p^2)/(1+q^2)\}}$ cannot be equal to p/q unless $p=q$, i.e. $x=1$. Thus all the points which correspond to rational values of x lie off the line, except the one point (1, 1). Again, if $p<q$, $\sqrt{\{(1+p^2)/(1+q^2)\}}>p/q$; if $p>q$, $\sqrt{\{(1+p^2)/(1+q^2)\}}<p/q$. Thus the points lie above the line $y=x$ if $0<x<1$, below if $x>1$. If p and q are large $\sqrt{\{(1+p^2)/(1+q^2)\}}$ is nearly equal to p/q. Near any value of x we can find any number of rational fractions with large numerators and denominators. Hence the graph contains a large number of points which crowd round the line $y=x$. Its general appearance (for positive values of x) is that of a line surrounded by a swarm of isolated points which gets denser and denser as the points approach the line.

The part of the graph which corresponds to negative values of x consists of the rest of the discontinuous line together with the reflections of all these isolated points in the axis of y. Thus to the left of the axis of y the swarm of points is not round $y=x$ but round $y=-x$, which is not itself part of the graph. See Fig. 19.

20. Graphical solution of Equations containing a single unknown quantity. Many equations can be expressed in the form

$$f(x)=\phi(x) \qquad \text{......................}(1),$$

where $f(x)$ and $\phi(x)$ are functions whose graphs are easy to draw. And it is obvious that if the curves

$$y=f(x), \quad y=\phi(x)$$

intersect in a point P whose abscissa is ξ, then ξ is a root of the equation (1).

Examples XVIII. 1. **The quadratic equation** $ax^2+2bx+c=0$. This may be solved graphically in a variety of ways. For instance we may draw the graphs of

$$y=ax+2b, \quad y=-c/x,$$

whose intersections give the roots, if any. Or we may take

$$y=x^2, \quad y=-(2bx+c)/a.$$

But the simplest method is probably to draw the circle

$$a(x^2+y^2)+2bx+c=0,$$

whose centre is $(-b/a, 0)$ and radius $\{\sqrt{(b^2-ac)}\}/a$. The abscissae of its intersections with the axis of x are the roots of the equation.

2. Solve by any of these methods

$$x^2+2x-3=0, \quad x^2-7x+4=0, \quad 3x^2+2x-2=0.$$

3. **The equation** $x^m+ax+b=0$. This may be solved by constructing the curves $y=x^m$, $y=-ax-b$.

4. Verify the following table for the number of real roots (if any) of

$$x^m+ax+b=0:$$

$$(a)\quad m \textit{ even} \begin{cases} b \text{ positive, } \textit{two or none}, \\ b \text{ negative, } \textit{two}; \end{cases}$$

$$(b)\quad m \textit{ odd} \begin{cases} a \text{ positive, } \textit{one}, \\ a \text{ negative, } \textit{three or one}. \end{cases}$$

Construct numerical examples to illustrate all possible cases.

5. Show that the equation $\tan x=ax+b$ has always an infinite number of real roots.

6. Determine the number of real roots of

$$\sin x=x, \quad \sin x=x/3, \quad \sin x=x/8, \quad \sin x=x/120.$$

7. Show that if a is small and positive (e.g. $a=\cdot 01$) the equation

$$x-a=\tfrac{1}{2}\pi \sin^2 x$$

has three real roots. Consider also the case in which a is small and negative. Explain how the number of roots varies as a varies.

21. Functions of two variables and their graphical representation. In § 9 we considered two variables connected by a relation. We may similarly consider *three* variables (x, y, and z) connected by a relation such that when the values of x and y are both given, the value or values of z are known. In this case we call z a *function of the two variables* x and y; x and y the *independent* variables, z the *dependent* variable; and we express this dependence of z upon x and y by writing

$$z=f(x, y).$$

The remarks of § 9 may all be applied, *mutatis mutandis*, to this more complicated case.

The method of representing such functions of two variables graphically is exactly the same in principle as in the case of functions of a single variable. We must take *three* axes OX, OY, OZ in space of three dimensions, each axis being perpendicular to the other two. The point (a, b, c) is the point whose distances from the planes YOZ, ZOX, XOY, measured parallel to OX, OY, OZ, are a, b, and c. Regard must of course be paid to sign, lengths measured in the directions OX, OY, OZ being regarded as positive. The definitions of *coordinates, axes, origin* are the same as before.

Now let $$z = f(x, y).$$
As x and y vary the point (x, y, z) will move in space. The aggregate of all these points is called the *locus* of the point (x, y, z) or the *graph* of the function $z = f(x, y)$. When the relation between x, y, and z which defines z can be expressed in an analytical formula this formula is called the *equation* of the locus.

22. **Equation of a plane.** It may be shown without difficulty that the coordinates of the point which divides PQ in a given ratio $\mu : \lambda$ are

$$\frac{\lambda a + \mu\alpha}{\lambda + \mu}, \quad \frac{\lambda b + \mu\beta}{\lambda + \mu}, \quad \frac{\lambda c + \mu\gamma}{\lambda + \mu},$$

where (a, b, c) are the coordinates of P and (α, β, γ) those of Q (cf. Ex. VIII. 3).

From this we can at once deduce the following important theorem: *the general equation of the first degree represents a plane.* For let

$$ax + by + cz + d = 0$$

be the equation; and let (x_1, y_1, z_1), (x_2, y_2, z_2) be two points P, Q on the graph of the function z (or the locus represented by the equation). Then

$$ax_1 + by_1 + cz_1 + d = 0,$$
$$ax_2 + by_2 + cz_2 + d = 0,$$

and so, multiplying by λ and μ, adding, and dividing by $\lambda + \mu$,

$$a\frac{\lambda x_1 + \mu x_2}{\lambda + \mu} + b\frac{\lambda y_1 + \mu y_2}{\lambda + \mu} + c\frac{\lambda z_1 + \mu z_2}{\lambda + \mu} + d = 0.$$

Thus the locus is such that if P and Q lie upon it, the point R which divides PQ in any ratio lies upon it. That is to say every point of the line PQ lies in the locus. The locus therefore satisfies Euclid's definition of a plane. Conversely *the equation of any plane is of the first degree.* For let (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3) be any three points on the plane. We can choose a, b, c, d so that

$$ax_1 + by_1 + cz_1 + d = 0,$$
$$ax_2 + by_2 + cz_2 + d = 0,$$
$$ax_3 + by_3 + cz_3 + d = 0.$$

We can therefore determine a locus represented by an equation of the type

$$ax + by + cz + d = 0,$$

which passes through the three points. But we have already seen that this locus is a plane; and it can obviously only be the original plane. The equation of that plane is therefore of the first degree.

Examples XIX. 1. Prove that if (a_1, b_1, c_1), (a_2, b_2, c_2), (a_3, b_3, c_3) are three points on a plane, the point whose coordinates are $\dfrac{\lambda a_1+\mu a_2+\nu a_3}{\lambda+\mu+\nu}$, etc. lies on the plane; and that it may, by choosing the ratios $\lambda:\mu:\nu$ appropriately, be made to coincide with any point in the plane.

2. Hence, by an argument similar to that of § 11, deduce that the general equation of the first degree represents a plane, and conversely.

3. The equation of the sphere whose centre is (a, b, c) and radius r is

$$(x-a)^2+(y-b)^2+(z-c)^2=r^2.$$

Conversely, this equation always represents a sphere.

4. Establish results for planes and spheres corresponding to those of Exs. IX. 3–9, 11–14.

23. Curves in a plane. We have hitherto used the notation

$$y=f(x) \qquad \text{..........................(1)}$$

to express functional dependence of y upon x. It is evident that this notation is most appropriate in the case in which y is expressed explicitly by means of some formula involving x alone, as when for example

$$y=x^2, \quad \sin x, \quad a\cos^2 x+b\sin^2 x.$$

We have however very often to deal with functional relations which cannot be or are most conveniently not expressed in this form. If, for example, $y^5-y-x=0$ or $x^5+y^5-ay=0$ it is known to be impossible to express y explicitly as a simple function of x. If

$$x^2+y^2+2gx+2fy+c=0$$

y can indeed be so expressed, viz. by the formula

$$y=-f\pm\sqrt{f^2-x^2-2gx-c};$$

but the functional dependence of y upon x is better and more simply expressed by the original equation.

It will be observed that in these two cases the functional relation is fully expressed *by equating a function of the two variables x and y to zero,* i.e. by means of an equation

$$f(x, y)=0 \qquad \text{..........................(2).}$$

We shall adopt this equation as the standard method of expressing the functional relation. It includes the equation (1) as a special case, since $y-f(x)$ is a special form of a function of x and y. We can then speak of the locus of the point (x, y) subject to $f(x, y)=0$, the graph of the function y defined by $f(x, y)=0$, the curve or locus $f(x, y)=0$, and the equation of this curve or locus.

There is another method of representing curves which is often useful. Suppose that x and y are both functions of a third variable t, which is to be regarded as essentially auxiliary and devoid of any particular geometrical significance. We may write

$$x=f(t), \quad y=F(t) \quad \text{.....................(3)}.$$

If t has any arbitrary value assigned to it, the value (or values) of x and of y are known. Each pair of such values defines a point (x, y). If we construct all the points which thus correspond to all the different values of t we obtain *the graph of the locus defined by the equations* (3). Suppose for example

$$x=a\cos t, \quad y=a\sin t.$$

Let t vary from 0 to 2π. Then it is easy to see that the point (x, y) describes the circle whose centre is the origin and radius is a. If t varies beyond these limits (x, y) describes the circle over and over again. We can in this case at once obtain a direct relation between x and y by squaring and adding: we find that $x^2+y^2=a^2$, t being now eliminated.

Examples XX. 1. The points of intersection of the two curves whose equations are $f(x, y)=0$, $\phi(x, y)=0$ are given by solving this pair of simultaneous equations.

2. Trace the curves $(x+y)^2=1$, $xy=1$, $x^2-y^2=1$.

3. The curve $f(x, y)+\lambda\phi(x, y)=0$ represents a curve passing through the points of intersection of $f=0$, $\phi=0$.

4. What loci are represented by

(α) $x=at+b$, $y=ct+d$, (β) $x/a=2t/(1+t^2)$, $y/a=(1-t^2)/(1+t^2)$,

when t varies through all real values?

24. Loci in space. In space of three dimensions there are two fundamentally different kinds of loci, of which the simplest examples are the plane and the straight line.

A particle which moves along a straight line has only *one degree of freedom.* Its direction of motion is fixed; if its velocity is given (velocity being regarded as a quantity capable of sign) its mode of motion is completely determined. Or again the position of a point on a line can be completely fixed by one measurement of position, e.g. by its distance from a fixed point on the line. If we take the line as our fundamental line L of Ch. I., the position of any of its points is determined by a *single coordinate* x. A particle which moves in a plane, on the other hand, has *two* degrees of freedom. In order to determine its mode of motion completely we require a knowledge of its component velocities in *two* different directions. Or again the position of a point on a plane requires the determination of *two* coordinates in order to fix it.

Now let us look at these loci from the point of view of their equations. The plane is represented by a single equation $\alpha x + \beta y + \gamma z + \delta = 0$. *Two* of the three coordinates y and z may be chosen arbitrarily, and the third is then fixed. The straight line on the other hand is the intersection of two planes. Let these two planes be

$$ax + by + cz + d = 0, \quad \alpha x + \beta y + \gamma z + \delta = 0 \quad\quad (1).$$

Then if *one* of the three coordinates is chosen arbitrarily, both of the others and the position of the point are fixed.

We can of course draw *any number* of planes through the line. Hence it might appear that the coordinates of the point on the line are subject to *more than two relations.* And so in fact they are, but the relations are not all independent. Any other plane through the line could be expressed in the form

$$ax + by + cz + d + \lambda(\alpha x + \beta y + \gamma z + \delta) = 0$$

and any equation of this type is a mere consequence of the equations (1).

The locus represented by a single equation

$$z = f(x, y)$$

is called a *surface.* It may or may not (in the obvious simple cases it will) satisfy our common-sense notion of what a surface should be.

The considerations of § 21 may evidently be generalised so as to give definitions of a function $f(x, y, z)$ of *three* variables (or of functions of any number of variables). And as in § 23 we

agreed to adopt $f(x, y)=0$ as the standard form of the equation of a plane curve, so now we shall agree to adopt

$$f(x, y, z)=0,$$

as the standard form of equation of a surface.

The locus represented by *two* equations of the form $z=f(x, y)$ or $f(x, y, z)=0$ is called a *curve*. Thus a *straight line* may be represented by two equations of the type $\alpha x+\beta y+\gamma z+\delta=0$. A *circle* in space may be regarded as the intersection of a sphere and a plane; it may therefore be represented by two equations of the forms

$$(x-a)^2+(y-b)^2+(z-c)^2=r^2, \quad \alpha x+\beta y+\gamma z+\delta=0.$$

Examples XXI. 1. What is represented by *three* equations of the type $f(x, y, z)=0$?

[Three equations in three variables are capable of solution (practically or theoretically). The solution consists of a finite or infinite number of isolated sets of values (x, y, z). The three equations therefore represent a number of isolated points.

Or we may regard the question thus. *Two* of the equations determine a *curve*, which meets the *surface* represented by the third equation in a number of points.]

2. Three linear equations represent a single point.

3. The equations of a curve differ from the equation of a surface in that their mode of expression is not unique, since either may be transformed by means of the other. Thus the curve $y=1$, $x^2+y^2+z^2=2$, (a circle) may also be represented by $y=1$, $x^2+z^2=1$.

4. What are the equations of a plane curve $f(x, y)=0$ in the plane XOY, when regarded as a curve in space? [$f(x, y)=0$, $z=0$.]

5. **Cylinders.** What is the meaning of a single equation $f(x, y)=0$, considered as a locus in space of three dimensions?

[All points on the surface satisfy $f(x, y)=0$ *whatever be the value* of z. The curve $f(x, y)=0$, $z=0$ is the curve in which the locus cuts the plane XOY. Draw the plane $z=a$, cutting ZOX, YOZ in OX', OY', and take OX', OY' as axes of coordinates in this plane (Fig. 20). Obviously $x'=x$, $y'=y$ and so $f(x', y')=0$. The curves in which the two planes $z=0$, $z=a$ cut the locus are therefore repetitions of the same plane curve: if one curve were moved a distance a parallel to the axis of z it would coincide with the other. The locus is the surface formed by drawing lines parallel to OZ through all points of the plane curve $f(x, y)=0$, $z=0$. Such a surface is called a *cylinder*.]

6. Interpret the equations: (*a*) $y=mx+c$, (*b*) $y=mx+c$, $z=a$, (*c*) $x^2+y^2=1$, (*d*) $x^2+y^2=1$, $z=a$, as loci in three-dimensional space.

7. Graphical representation of a surface on a plane. Contour Maps. It might seem to be impossible to adequately represent a surface by a drawing on a plane; and so indeed it is: but a very fair notion of the nature of the surface may often be obtained as follows. Let the equation of the surface be $z=f(x, y)$.

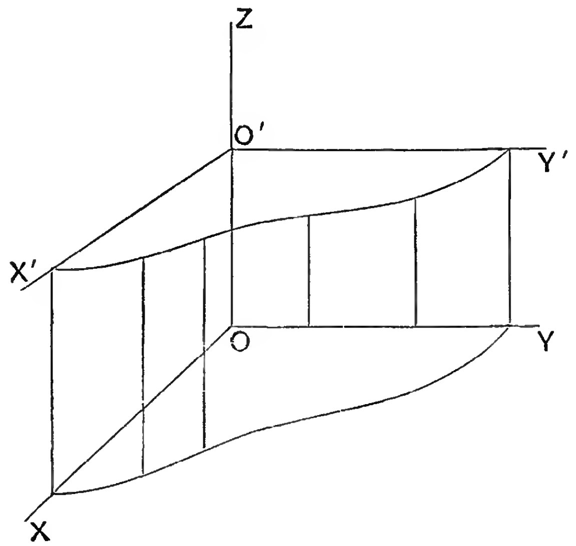


FIG. 20.

If we give z a particular value a, we have an equation $a=f(x, y)$, which we may regard as determining a plane curve on the paper. We trace this curve and mark it (a). Actually the curve (a) is the projection on the plane XOY of the section of the surface by the plane $z=a$ (Fig. 20). We do this for all values of a (practically, of course, for a selection of values of a). We obtain some such figure as is shown in Fig. 21. It will at once suggest a contoured Ordnance Survey map: and in fact this is the principle on which such maps are constructed. The contour line 1000 is the projection on the plane of the sea level of the section of the surface of the land by the plane parallel to the plane of the sea level and 1000 ft. above it*.

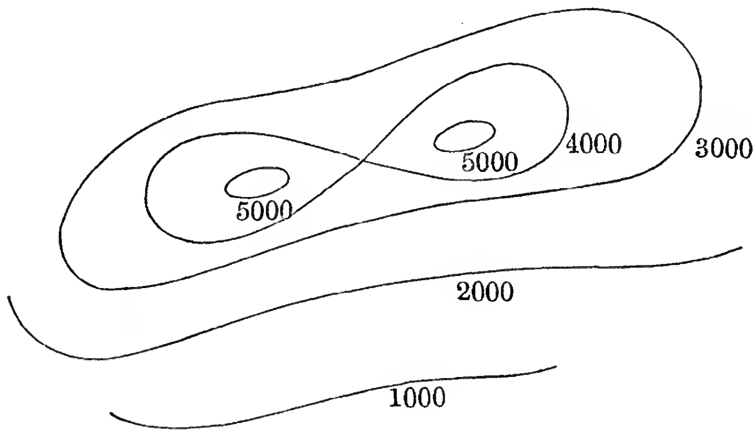


FIG. 21.

* We may assume here that the effects of the earth's curvature may be neglected.

8. Draw a series of contour lines to illustrate the form of the surface $2z=3xy$.

9. **Right circular cones.** Take the origin of coordinates at the vertex of the cone and the axis of z along the axis of the cone (Figs. 22, 23). Let a be the semi-vertical angle of the cone, P any point on it, C and P' its projections

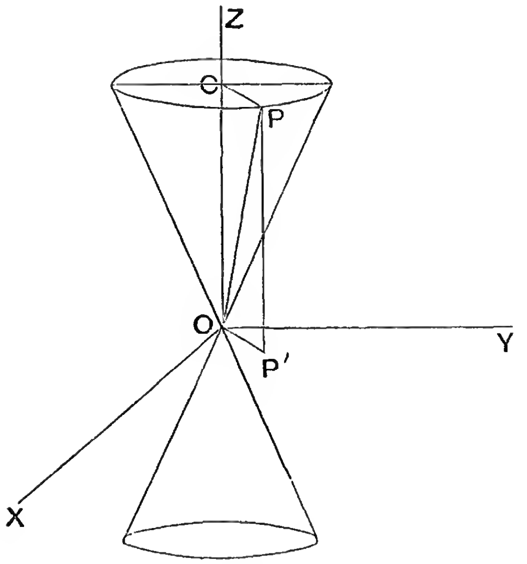


FIG. 22.

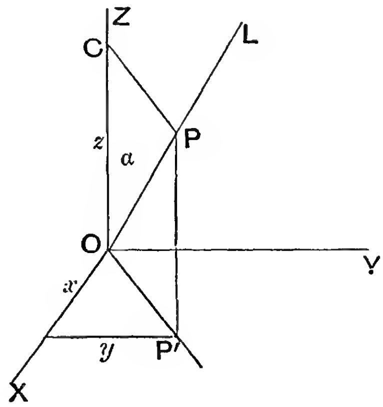


FIG. 23.

on the axis OZ and the plane XOY. Then if x, y, z are the coordinates of P, we have $x^2+y^2=OP'^2=CP^2=OC^2\tan^2 a=z^2\tan^2 a$. The equation of the cone (which must be regarded as extending both ways from its vertex) is therefore $x^2+y^2-z^2\tan^2 a=0$.

10. **Surfaces of revolution in general.** We notice that the cone of Ex. 9 cuts ZOX in the lines $x=\pm z\tan a$, which may be combined in the equation $x^2=z^2\tan^2 a$. That is to say, the equation of the surface generated by the revolution of the curve $y=0$, $x^2=z^2\tan^2 a$ round the axis of z is derived from the second of these equations by changing x^2 into x^2+y^2.

Show generally that the equation of the surface generated by the revolution of the curve $y=0$, $x=f(z)$, round the axis of z, is $\sqrt{(x^2+y^2)}=f(z)$, or

$$x^2+y^2=\{f(z)\}^2.$$

Verify in the case of (1) the line $y=0$, $x=1$; (2) the circle $y=0$, $x^2+z^2=1$.

11. **Cones in general.** A surface formed by straight lines passing through a fixed point is called a *cone*: the point is called the *vertex*. A particular case is given by the right circular cone of Ex. 9. Show that the equation of a cone whose vertex is O is of the form

$$f\left(\frac{z}{x}, \frac{z}{y}\right)=0,$$

and that any equation of this form represents a cone.

[If (x, y, z) lies on the cone, so must $(\lambda x, \lambda y, \lambda z)$, for any value of λ.]

12. **Ruled surfaces.** Cylinders and cones are special cases of *surfaces composed of straight lines.* Such surfaces are called *ruled surfaces.*

The two equations

$$\left.\begin{aligned} x &= az + b \\ y &= cz + d \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

represent the intersection of two planes, i.e. a straight line. Now suppose that a, b, c, d instead of being fixed, are *functions of an auxiliary variable* t. For any particular value of t the equations (1) give a line. As t varies this line moves, and generates a surface, whose equation may be found by eliminating t between the two equations (1). For instance, in Fig. 23 the line OL is inclined at a fixed angle α to OZ. PP' is perpendicular to the plane XOY and $XOP' = t$. The equations of the line are

$$\left.\begin{aligned} x &= z \tan \alpha \cos t \\ y &= z \tan \alpha \sin t \end{aligned}\right\}.$$

As t varies the line turns round OZ and generates the cone $x^2 + y^2 = z^2 \tan^2 \alpha$.

Another simple example of a ruled surface may be constructed as follows. Take two sections of a right circular cylinder perpendicular to the axis and at a distance l apart (Fig. 24 a). We can imagine the surface of the cylinder to be made up of a number of thin parallel rigid rods of length l, such as PQ, the ends of the rods being fastened to two circular rods of radius a.

Now let us take a third circular rod of the same radius and place it round the surface of the cylinder at a distance h from one of the first two rods (Fig. 24 a). Unfasten the end Q of the rod PQ and turn PQ about P until Q can be fastened to the third circular rod in the position Q'. The angle $qOQ' = \alpha$ in the figure is evidently given by

$$l^2 - h^2 = qQ'^2 = (2a \sin \tfrac{1}{2}\alpha)^2.$$

Let all the other rods of which the cylinder was composed be treated in the same way. We obtain a ruled surface whose form is indicated in Fig. 24 b. It is entirely built up of straight lines; but the surface is curved everywhere, and is in general shape not unlike certain forms of table-napkin rings (Fig. 24c).

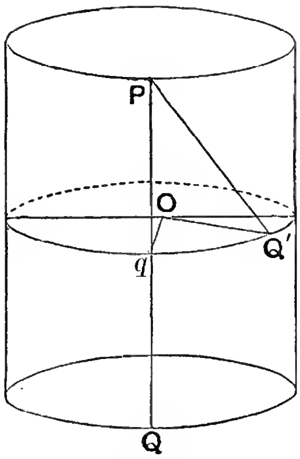


FIG. 24 a.

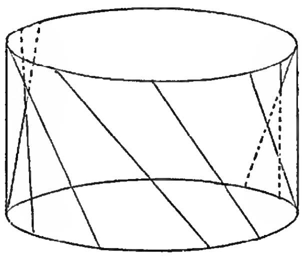


FIG. 24 b.

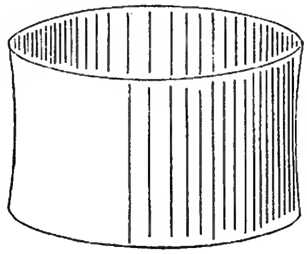


FIG. 24 c.

MISCELLANEOUS EXAMPLES ON CHAPTER II.

1. If $y=f(x)=(ax+b)/(cx-a)$, show that $x=f(y)$.

2. If $f(x)=f(-x)$ for all values of x, $f(x)$ is called an *even* function. If $f(x)=-f(-x)$ it is called an *odd* function. Show that any function of x, defined for all values of x, is the sum of an even and an odd function of x. [Use the identity $f(x)=\frac{1}{2}\{f(x)+f(-x)\}+\frac{1}{2}\{f(x)-f(-x)\}$.]

3. Find all the values of x for which $y=\dfrac{x-\sqrt{2}}{\sqrt{3}-x}$ has a rational value.

4. Draw the graphs of the functions

$$3\sin x+4\cos x, \quad \sin\left(\frac{\pi}{\sqrt{2}}\sin x\right). \qquad \text{(}\textit{Math. Trip.}\text{ 1896.)}$$

5. Draw the graphs of the functions

$$\sin x\,(a\cos^2 x+b\sin^2 x), \quad \frac{\sin x}{x}(a\cos^2 x+b\sin^2 x), \quad \left(\frac{\sin x}{x}\right)^2.$$

6. Draw graphs of the functions

(i) $\arccos(2x^2-1)-2\arccos x$,

(ii) $\arctan\dfrac{a+x}{1-ax}-\arctan a-\arctan x$,

where the symbols $\arccos a$, $\arctan a$ denote, for any value of a, the least positive (or zero) angle, whose cosine or tangent is a.

7. Verify the following method of constructing the graph of $f\{\phi(x)\}$ by means of the line $y=x$ and the graphs of $f(x)$ and $\phi(x)$: take $OA=x$ along OX, draw AB parallel to OY to meet $y=\phi(x)$ in B, BC parallel to OX to meet $y=x$ in C, CD parallel to OY to meet $y=f(x)$ in D, and DP parallel to OX to meet AB in P: then P is a point on the graph required.

8. Show that the roots of $x^3+px+q=0$ are the abscissae of the points of intersection (other than the origin) of the parabola $y=x^2$ and the circle

$$x^2+y^2+(p-1)y+qx=0.$$

9. The roots of $x^4+nx^3+px^2+qx+r=0$ are the abscissae of the points of intersection of the parabola $x^2=y-\frac{1}{2}nx$, and the circle

$$x^2+y^2+(\tfrac{1}{8}n^2-\tfrac{1}{2}pn+\tfrac{1}{2}n+q)\,x+(p-1-\tfrac{1}{4}n^2)\,y+r=0.$$

10. Discuss the graphical solution of the equation

$$x^m+ax^2+bx+c=0$$

by means of the curves $y=x^m$, $y=-ax^2-bx-c$. Draw up a table of the various possible numbers of real roots.

11. Show that the equation

$$2x=(2n+1)\pi(1-\cos x)$$

where n is a positive integer, has $2n+3$ real roots and no more, roughly indicating their localities. (*Math. Trip.* 1896.)

12. Discuss the number and value of the real roots of the equations

(1) $\cot x+x-\frac{3}{2}\pi=0$, (2) $x^2+\sin^2 x=1$, (3) $\tan x=2x/(1+x^2)$,

(4) $\sin x-x+\frac{1}{6}x^3=0$, (5) $(1-\cos x)\tan a-x+\sin x=0$.

13. Determine a polynomial of the 5th degree which has, for $x=-1$, $-\frac{1}{2}$, 0, $\frac{1}{2}$, 1 the values 3, 7, 2, 0, 4.

14. The polynomial of the second degree which assumes, when $x=a, b, c$, the values α, β, γ, is

$$\alpha\frac{(x-b)(x-c)}{(a-b)(a-c)}+\beta\frac{(x-c)(x-a)}{(b-c)(b-a)}+\gamma\frac{(x-a)(x-b)}{(c-a)(c-b)}.$$

Give a similar formula for the polynomial of the $(n-1)$-th degree which assumes, when $x=a_1, a_2, \ldots a_n$, the values $\alpha_1, \alpha_2, \ldots \alpha_n$.

15. If x is a rational function of y, and y is a rational function of x, show that $Axy+Bx+Cy+D=0$.

16. If y is a rational function of x, with rational coefficients, then y has a rational value for all rational values of x.

17. If y is an algebraical function of x, x is an algebraical function of y.

18. Verify that for values of x between 0 and 1 the equation

$$\cos\tfrac{1}{2}\pi x=1-\frac{x^2}{x+(x-1)\sqrt{\left(\dfrac{2-x}{3}\right)}}$$

is approximately true. [Take $x=0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}, 1$, and use tables. For which of these values is the formula exact?]

19. The equation

$$a_0x^n+na_1x^{n-1}y+\frac{n(n-1)}{1\,.\,2}a_2x^{n-2}y^2+\ldots+a_ny^n=0$$

represents n straight lines through the origin.

20. Show that the line $Ax+By+C=0$ and the two lines represented by

$$(Ax+By)^2-3(Ay-Bx)^2=0$$

form the sides of an equilateral triangle. (*Math. Trip.* 1906.)

21. The equation of the circle described on the line joining (x, y) and (x', y') as diameter is $(x-x')(x-x'')+(y-y')(y-y'')=0$.

22. The general equation of a circle through (x', y') and (x'', y'') may be expressed in either of the forms

(1) $(x-x')(y-y'')-(x-x'')(y-y')$
$$=\{(x-x')(x-x'')+(y-y')(y-y'')\}\tan\alpha,$$

$$(2) \quad (x-x')\left(x-x''+\frac{\lambda}{x''-x'}\right)+(y-y')\left(y-y''-\frac{\lambda}{y''-y'}\right)=0.$$

Here a is the angle contained in one of the segments of the circle. Express λ in terms of a.

23. The general equation of a circle cutting $x^2+y^2+2\lambda x+c=0$ orthogonally, for all values of λ, is $x^2+y^2+2\mu y-c=0$. Sketch the two sets of circles.

24. The general equation of all circles cutting at right angles the two circles $x^2+y^2-2a_1x-2b_1y+c_1=0$, $x^2+y^2-2a_2x-2b_2y+c_2=0$ is

$$\begin{vmatrix} x^2+y^2 & x & y \\ c_1 & a_1 & b_1 \\ c_2 & a_2 & b_2 \end{vmatrix} + \lambda \begin{vmatrix} x & y & 1 \\ a_1 & b_1 & 1 \\ a_2 & b_2 & 1 \end{vmatrix} = 0.$$

(*Math. Trip.* 1906.)

25. Show that the intersection of the two circular cylinders $x^2+y^2=1$, $x^2+z^2=1$, consists of two plane curves. Give a sketch of the cylinders and their line of intersection.

26. **Sections of a right circular cone by a plane.** Show that the cone $x^2+y^2=z^2\tan^2 a$ and the plane $z=x\tan\theta+c$ intersect in a curve whose projection on the plane XOY is $x^2+y^2=(x\tan\theta+c)^2\tan^2 a$.

Taking axes $O'\xi$, $O'\eta$ in the plane of section, O' being on OZ and $O'\eta$ parallel to OY, show that the equation of the curve of section is

$$\xi^2\cos^2\theta+\eta^2=(\xi\sin\theta+c)^2\tan^2 a \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1).$$

Show that this curve consists of a single closed branch, a single infinite branch, or two infinite branches, according as $\theta \lesseqgtr \frac{1}{2}\pi - a$, and that in any case it is symmetrical about $O'\xi$.

27. Show that the equation (1) of the last example may be expressed in the form

$$(\xi-\gamma)^2+\eta^2=e^2(\xi-\kappa)^2,$$

where
$$e=\sin\theta\sec a, \quad \gamma=-\frac{c\sin a\sin\theta}{\sin a\pm\cos\theta},$$

and
$$\kappa=-\frac{c\sin a}{\sin\theta}\,\frac{1\pm\sin a\cos\theta}{\sin a\pm\cos\theta};$$

unless $\theta=\frac{1}{2}\pi-a$, in which case

$$e=1, \quad \gamma=-\tfrac{1}{2}c\cos a, \quad \kappa=-\tfrac{1}{2}c\sec a\,(1+\sin^2 a).$$

28. Deduce that the section is a curve such that the distance of any point on the curve from a fixed point $(\gamma, 0)$ is e times its distance from a fixed line $\xi-\kappa=0$, i.e. that the curve is a conic, having the focus and directrix property which is usually adopted as the definition of a conic in books on Conic Sections. The conic is an ellipse, parabola, or hyperbola, according as $e \lesseqgtr 1$; and, except in the special cases when $e=1$ or $e=0$, has two foci $(\gamma, 0)$ and two corresponding directrices.

29. Let A, A' be the vertices of the conic, i.e. the points where the conic cuts the axis of symmetry $O'\xi$, S, S' the foci, and K, K' the points where the directrices cut $O'\xi$. Show that A, A' are given by $\xi = \mp c \sin \alpha / \cos(\alpha \mp \theta)$, that A lies between S and K and A' between S' and K', and that the length of the 'major axis' AA' is $2c \sin \alpha \cos \alpha \cos \theta / (\cos^2 \theta - \sin^2 \alpha)$.

30. Show further that if we take axes parallel to $O'\xi$, $O'\eta$ through C, the middle point of AA', the equation of the curve becomes of the form

$$(x^2/a^2) + (y^2/b^2) = 1,$$

where $a = \frac{1}{2}AA'$ and $b = a\sqrt{(1-e^2)}$, or

$$(x^2/a^2) - (y^2/b^2) = 1,$$

where $a = \frac{1}{2}AA'$ and $b = a\sqrt{(e^2-1)}$, according as $e \lessgtr 1$. Sketch the forms of the curves.

31. In the case when $e=1$, show that the one point A where the curve cuts $O'\xi$ is given by $\xi = -\frac{1}{2}c \sec \alpha$, and that by taking axes through this point we can reduce the equation of the curve to the form $y^2 = 4ax$, where $a = \frac{1}{2}c \sin \alpha \tan \alpha$.

[For an account of the simplest properties of the conic sections, deducible from the equations $(x^2/a^2) \pm (y^2/b^2) = 1$ or $y^2 = 4ax$, we must refer to treatises dealing specially with this subject.]

32. Show that the most general equation of the second degree, viz.

$$ax^2 + 2hxy + by^2 + 2gx + 2fy + c = 0$$

represents a conic. [It is this property which accounts for the importance of the conic sections.]

33. Show that the equation represents an ellipse, parabola, or hyperbola, according as $h^2 \lesseqgtr ab$.

34. The equation

$$ax^2 + 2hxy + by^2 + 2(gx+fy)(lx+my) + c(lx+my)^2 = 0$$

represents the two lines joining the origin to the points in which the line $lx+my=1$ cuts the conic $ax^2+2hxy+by^2+2gx+2fy+c=0$.

35. Show directly that if the cylinder $x^2+y^2=1$ is cut by a plane neither parallel nor perpendicular to its axis, the intersection is a curve possessing the focus and directrix property of a conic.

36. The curve

$$\lambda f(x, y)\,\phi(x, y) + \mu F(x, y)\,\Phi(x, y) = 0$$

passes through all points of intersection of $f=0$ and $F=0$, of $f=0$ and $\Phi=0$, of $\phi=0$ and $F=0$, and of $\phi=0$ and $\Phi=0$.

37. If $\lambda_r = a_r x + b_r y + c_r$, the equation $\lambda L_1 L_3 + \mu L_2 L_4 = 0$ is the general equation of a conic circumscribing the quadrangle formed by the four lines $L_1=0$, $L_2=0$, $L_3=0$, $L_4=0$ taken in order.

38. Determine the equation of the surface generated when the circle

$$y=0, \quad (x-a)^2+z^2=1$$

rotates round the axis of z. Sketch the form of the surface for different values of a.

39. What is the form of the graph of the functions

$$z=[x]+[y], \quad z=x+y-[x]-[y]\,?$$

40. What is the form of the graph of the functions $z=\sin x+\sin y$, $z=\sin x \sin y$, $z=\sin xy$, $z=\sin(x^2+y^2)$?

41. **Geometrical Constructions for irrational numbers.** In Chapter I. we indicated one or two simple geometrical constructions for a length equal to $\sqrt{2}$, starting from a given unit length. We also showed how to construct the roots of any quadratic equation $ax^2+2bx+c=0$, it being supposed that we can construct lines whose lengths are equal to any of the ratios of the quantities a, b, c. All these constructions were what may be called Euclidean constructions; they depended on the ruler and compass only.

It is fairly obvious that any irrational expression, however complicated, can be constructed by means of these methods, *provided it only contains square roots.* Thus

$$\sqrt[4]{\left\{\sqrt{\left(\frac{17+3\sqrt{11}}{17-3\sqrt{11}}\right)}-\sqrt{\left(\frac{17-3\sqrt{11}}{17+3\sqrt{11}}\right)}\right\}}$$

is a case in point. This contains a fourth root, but this is of course the square root of a square root. We should begin by constructing $\sqrt{11}$, e.g. as the mean between 1 and 11: then $17+3\sqrt{11}$, and so on. Or these two mixed surds might be constructed directly as the roots of $x^2-34x+190=0$.

Conversely, *only* irrationals of this kind *can* be constructed by Euclidean methods. Starting from a unit length we can construct any *rational* length. And hence we can construct the line $ax+by+c=0$, or the circle

$$(x-\alpha)^2+(y-\beta)^2=r^2, \quad (\text{or } x^2+y^2+2gx+2fy+d=0)$$

provided the constants which occur in these equations are rational.

Now in any Euclidean construction, each new point introduced into the figure is determined as the intersection of two lines or circles, or a line and a circle. But if the coefficients are rational, such a pair of equations as

$$ax+by+c=0, \quad x^2+y^2+2gx+2fy+d=0$$

give, on solution, values of x and y of the form $m+n\sqrt{p}$, where m, n, p are rational: for if we substitute for x in terms of y in the second equation we obtain a quadratic in y with rational coefficients. Hence the coordinates of all points obtained by means of lines and circles with rational coefficients are expressible by rational numbers and quadratic surds. And so the same is true of the distance $\sqrt{\{(x_1-x_2)^2+(y_1-y_2)^2\}}$ between any two points so obtained.

With the irrational distances thus constructed we may proceed to construct

a number of lines and circles whose coefficients may now themselves involve quadratic surds. It is evident, however, that by the use of such lines and circles we can still only construct lengths expressible by square roots only, though our surd expressions may now be of a more complicated form. And it is clear that this remains true however far we may go. Hence *Euclidean methods will construct any surd expression involving square roots, and no others.* In particular they will not construct $\sqrt[3]{2}$, i.e. they will not solve the problem of *the duplication of the cube*, which was one of the famous problems of antiquity.

42. **Approximate quadrature of the circle.** Let O be the centre of a circle of radius R. On the tangent at A take $AP = \frac{11}{5}R$ and $AQ = \frac{13}{5}R$, in the same direction. On AO take $AN = OP$ and draw NM parallel to OQ and cutting AP in M. Show that

$$AM = \tfrac{13}{25}\sqrt{146}\,.\,R,$$

and that to take AM as being equal to the circumference of the circle would lead to a value of π correct to five places of decimals.

If R is the earth's radius, the error in supposing AM to be its circumference is less than 11 yards.

43. Show that the only lengths which can be constructed *with the ruler only*, starting from a given unit length, are *rational* lengths.

44. **Constructions for $\sqrt[3]{2}$.** O is the vertex and S the focus of the parabola $y^2 = 4x$, and P is one of its points of intersection with the parabola $x^2 = 2y$. Show that OP meets the latus rectum of the first parabola in a point Q such that $SQ = \sqrt[3]{2}$.

45. Take a circle of unit diameter, a diameter OA and the tangent at A. Draw a chord OBC cutting the circle at B and the tangent at C. On this line take $OM = BC$. Taking O as origin and OA as axis of x, show that the locus of M is the curve

$$(x^2 + y^2)\,x - y^2 = 0$$

(the *Cissoid of Diocles*). Sketch the curve. Take along the axis of y a length $OD = 2$. Let AD cut the curve in P and OP cut the tangent at A in Q. Show that $AQ = \sqrt[3]{2}$.

CHAPTER III.

COMPLEX NUMBERS.

25. Displacements along a line and in a plane. The 'real number' x, with which we have been concerned in the two preceding chapters, may be regarded from a considerable number of different points of view. It may be regarded as a *pure number*, destitute of geometrical significance, or a geometrical significance may be attached to it in at least three different ways. It may be regarded as *the measure of a length*, viz. the length A_0P along the line L of Chap. I. It may be regarded as *the mark of a point*, viz. the point P whose distance from A_0 is x. Or it may be regarded as *the measure of a displacement* or *change of position* on the line L. It is on this last point of view that we shall now concentrate our attention.

Let a small particle be placed at P on the line L and then displaced to Q. We shall call the displacement or change of position which is needed to transfer the particle from P to Q *the displacement* $\overline{PQ}$. To completely specify a displacement three things are needed, its *magnitude*, its *sense* (forwards or backwards along the line), and what may be called its *point of application*, i.e. the original position P of the particle. But, when we are thinking merely of the change of position produced by the displacement, it is natural to disregard the point of application and to consider all displacements as equivalent whose lengths and senses are the same. Then the displacement is completely specified by the length $PQ = x$, the sense of the displacement being fixed by the sign of x. We may therefore, without ambiguity, speak of *the displacement* $[x]$, and we may write

$$\overline{PQ} = [x].$$

We use the square bracket to distinguish the displacement $[x]$ from the length or number x*. If the coordinate of P is a, that of Q will be $a+x$; the displacement $[x]$ therefore transfers a particle from the point x to the point $a+x$.

We come now to consider *displacements in a plane*. We may define the displacement $\overline{PQ}$ as before. But now more data are required in order to specify it completely. We require to know: (i) the *magnitude* of the displacement, i.e. the length of the straight line PQ; (ii) the *direction* of the displacement, which is determined by the angle which PQ makes with some fixed line in the plane; (iii) the *sense* of the displacement; and (iv) its *point of application*. Of these requirements we may disregard the fourth, if we consider two displacements as equivalent if they are the same in magnitude, direction, and sense. In other words, if PQ and RS are equal and parallel, and the sense of motion from P to Q is the same as that of motion from R to S, we regard the displacements $\overline{PQ}$ and $\overline{RS}$ as equivalent, and write

$$\overline{PQ} = \overline{RS}.$$

Now let us take any pair of coordinate axes in the plane

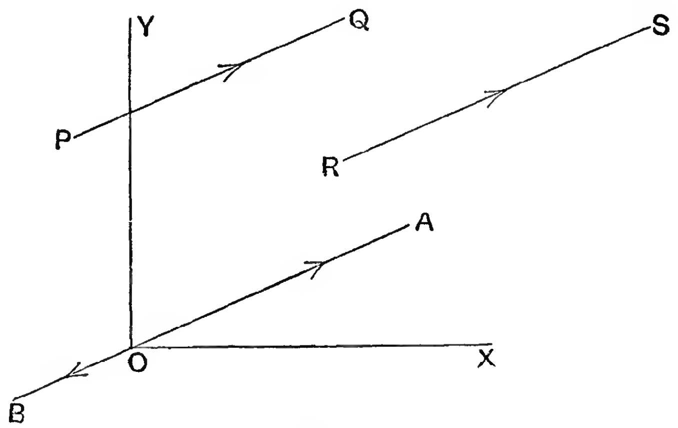


FIG. 25.

(such as OX, OY in Fig. 25). Draw a line OA equal and parallel

* Strictly speaking we ought, by some similar difference of notation, to distinguish the actual length x from the number x which measures it. The reader will perhaps be inclined to consider such distinctions futile and pedantic. But increasing experience of mathematics will reveal to him the great importance of distinguishing clearly between things which, however intimately connected, are *not the same*. If cricket were a mathematical science it would be very important to distinguish between the *motion* of the batsman between the wickets, the *run* which he scores, and the *mark* which is put down in the score-book.

5—2

to PQ, the sense of motion from O to A being the same as that from P to Q. Then $\overline{PQ}$ and $\overline{OA}$ are equivalent displacements. Let x and y be the coordinates of A. Then it is evident that $\overline{OA}$ is completely specified if x and y are given. We call $\overline{OA}$ *the displacement* $[x, y]$ and write

$$\overline{OA} = \overline{PQ} = \overline{RS} = [x, y].$$

26. Equivalence of displacements. Multiplication of displacements by numbers. If ξ and η are the coordinates of P, ξ' and η' those of Q, it is evident that

$$x = \xi' - \xi, \quad y = \eta' - \eta.$$

The displacement from (ξ, η) to (ξ', η') is therefore

$$[\xi' - \xi, \ \eta' - \eta].$$

It is evident that two displacements $[x, y]$, $[x', y']$ are equivalent if, and only if, $x = x'$, $y = y'$. Thus $[x, y] = [x', y']$ if

$$x = x', \quad y = y' \quad \text{.......................(1).}$$

The reverse displacement $\overline{QP}$ would be $[\xi - \xi', \ \eta - \eta']$, and it is natural to agree that

$$[\xi - \xi', \ \eta - \eta'] = -[\xi' - \xi, \ \eta' - \eta],$$
$$\overline{QP} = -\overline{PQ},$$

these equations being really definitions of the meaning of the symbols $-[\xi' - \xi, \ \eta' - \eta]$, $-\overline{PQ}$.

Having thus agreed that

$$-[x, y] = [-x, -y],$$

it is natural to agree further that

$$\alpha[x, y] = [\alpha x, \alpha y] \quad \text{.....................(2)}$$

where α is any real number, positive or negative. Thus (Fig. 25) if $OB = -\frac{1}{2}OA$,

$$\overline{OB} = -\tfrac{1}{2}\overline{OA} = -\tfrac{1}{2}[x, y] = [-\tfrac{1}{2}x, -\tfrac{1}{2}y].$$

The equations (1) and (2) define the first two important ideas connected with displacements, viz. *equivalence* of displacements, and *multiplication of displacements by numbers*.

27. Addition of displacements. We have not yet given any definition which enables us to attach any meaning to the expressions

$$\overline{PQ} + \overline{P'Q'},\quad [x, y] + [x', y'].$$

Common sense at once suggests that we should define the sum of two displacements as the displacement which is the result

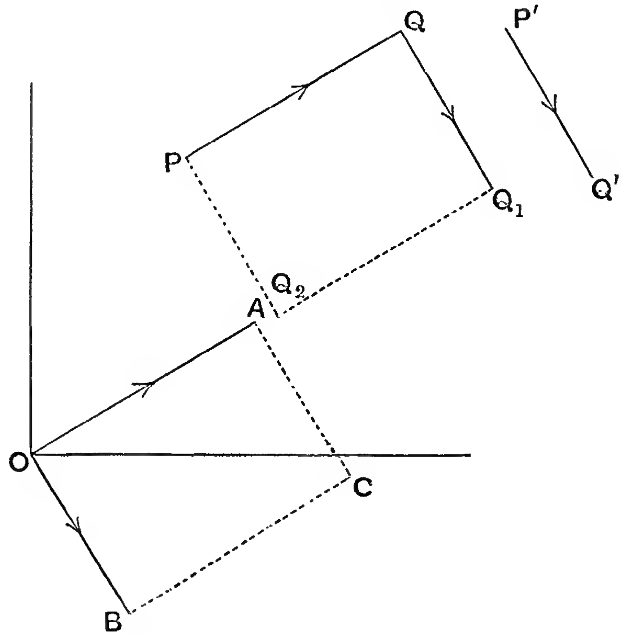


Fig. 26.

of the successive application of the two given displacements. In other words, it suggests that if QQ_1 be drawn equal and parallel to $P'Q'$, so that the result of successive displacements $\overline{PQ}$, $\overline{P'Q'}$ on a particle at P is to transfer it first to Q and then to Q_1, we should define the sum of $\overline{PQ}$ and $\overline{P'Q'}$ as being $\overline{PQ_1}$. Or, if we draw OB equal and parallel to $P'Q'$, and complete the parallelogram $OACB$,

$$\overline{PQ} + \overline{P'Q'} = \overline{OA} + \overline{OB} = \overline{OC} = \overline{PQ_1}.$$

Let us consider the consequences of adopting this definition. If the coordinates of B are x', y', those of the middle point of AB are $\frac{1}{2}(x+x')$, $\frac{1}{2}(y+y')$, and those of C are $x+x'$, $y+y'$. Hence

$$[x, y] + [x', y'] = [x + x', y + y'] \quad \text{............(3)},$$

which may be regarded as the symbolic definition of *addition of displacements.* We observe that

$$\begin{aligned} [x', y'] + [x, y] &= [x' + x, y' + y] \\ &= [x + x', y + y'] = [x, y] + [x', y']. \end{aligned}$$

In other words, *addition of displacements obeys the commutative law* expressed in ordinary algebra by the equation $a+b=b+a$.

Looked at geometrically, this expresses the obvious fact that if we move from P *first* through a distance PQ_2 equal and parallel to $P'Q'$, and then through a distance equal and parallel to PQ, we shall arrive at the same point Q_1 as before. Again, since

$$[x, y]+[x, y]=[2x, 2y]=2\,[x, y],$$

our definition of addition agrees with that previously adopted for multiplication by a number.

In particular

$$[x, y]=[x, 0]+[0, y] \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4).$$

Here $[x, 0]$ denotes a displacement through a distance x in a direction parallel to OX. It is in fact what we previously denoted by $[x]$, when we were considering only displacements along a line. We call $[\cancel{y}\,x, 0]$ and $[0, y]$ the *components* of $[x, y]$, and $[x, y]$ their *resultant.*

When we have once defined addition of two displacements there is no further difficulty in the way of defining addition of any number. Thus (by definition)

$$[x, y]+[x', y']+[x'', y'']=([x, y]+[x', y'])+[x'', y'']$$
$$=[x+x', y+y']+[x'', y'']=[x+x'+x'', y+y'+y''].$$

We define *subtraction* of displacements by the equation

$$[x, y]-[x', y']=[x, y]+(-[x', y']) \quad\ldots\ldots\ldots\ldots(5),$$

which is the same thing as $[x, y]+[-x', -y']$ or as $[x-x', y-y']$.

In particular

$$[x, y]-[x, y]=[0, 0].$$

The displacement $[0, 0]$ *leaves the particle where it was*; it is the *zero displacement*, and we agree to write $[0, 0]=0$.

Examples XXII. 1. Addition of displacements, and multiplication of displacements by numbers, obey all the ordinary laws of algebra, expressed by the equations,

(i) $\alpha\,[\beta x, \beta y]=\beta\,[\alpha x, \alpha y]=[\alpha\beta x, \alpha\beta y]$,

(ii) $([x, y]+[x', y'])+[x'', y'']=[x, y]+([x', y']+[x'', y''])$,

(iii) $[x, y]+[x', y']=[x', y']+[x, y]$,

(iv) $(\alpha+\beta)\,[x, y]=\alpha\,[x, y]+\beta\,[x, y]$,

(v) $\alpha\,\{[x, y]+[x', y']\}=\alpha\,[x, y]+\alpha\,[x', y']$.

[We have already proved (iii). The remaining equations follow with equal ease from the definitions. The reader should in each case consider the geometrical significance of the equation, as we did above in the case of (iii).]

2. If M is the middle point of PQ, $\overline{OM}=\frac{1}{2}(\overline{OP}+\overline{OQ})$. More generally if M divides PQ in the ratio $\mu:\lambda$

$$\overline{OM}=\frac{\lambda}{\lambda+\mu}\overline{OP}+\frac{\mu}{\lambda+\mu}\overline{OQ}.$$

3. If G is the centre of mass of equal particles at $P_1, P_2, \dots, P_n$

$$\overline{OG}=(\overline{OP}_1+\overline{OP}_2+\dots+\overline{OP}_n)/n.$$

4. If P, Q, R are collinear points in the plane, it is possible to find real numbers α, β, γ, not all zero, and such that

$$\alpha\,.\,\overline{OP}+\beta\,.\,\overline{OQ}+\gamma\,.\,\overline{OR}=0\,;$$

and conversely. [This is really only another way of stating Ex. 2.]

5. If $\overline{AB}$ and $\overline{AD}$ are two displacements not in the same straight line, and

$$\alpha\,.\,\overline{AB}+\beta\,.\,\overline{AD}=\gamma\,.\,\overline{AB}+\delta\,.\,\overline{AD},$$

then $\alpha=\gamma$ and $\beta=\delta$.

[Take $AB_1=\alpha\,.\,AB$, $AD_1=\beta\,.\,AD$. Complete the parallelogram $AB_1P_1D_1$. Then $\overline{AP}_1=\alpha\,.\,\overline{AB}+\beta\,.\,\overline{AD}$. It is evident that $\overline{AP}_1$ can only be expressed in this form in one way, whence the theorem follows.]

6. $ABCD$ is a parallelogram. Through Q, a point inside the parallelogram, RQS and TQU are drawn parallel to the sides. Show that RU, TS intersect on AC.

[Let the ratios $AT:AB$, $AR:AD$ be denoted by α, β. Then

$$\overline{AT}=\alpha\,.\,\overline{AB},\quad \overline{AR}=\beta\,.\,\overline{AD},$$

$$\overline{AU}=\alpha\,.\,\overline{AB}+\overline{AD},\quad \overline{AS}=\overline{AB}+\beta\,.\,\overline{AD}.$$

Let RU meet AC in P. Then, since R, U, P are collinear

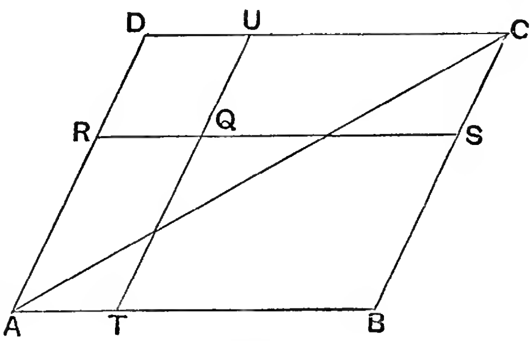


Fig. 27.

$$\overline{AP}=\frac{\lambda}{\lambda+\mu}\overline{AR}+\frac{\mu}{\lambda+\mu}\overline{AU},$$

where μ/λ is the ratio in which P divides RU. That is to say

$$\overline{AP}=\frac{\alpha\mu}{\lambda+\mu}\overline{AB}+\frac{\beta\lambda+\mu}{\lambda+\mu}\overline{AD}.$$

But since P lies on AC, $\overline{AP}$ is a numerical multiple of $\overline{AC}$; say

$$\overline{AP}=k\,.\,\overline{AC}=k\,.\,\overline{AB}+k\,.\,\overline{AD}.$$

Hence (Ex. 5) $\alpha\mu=\beta\lambda+\mu=(\lambda+\mu)k$, from which we deduce

$$k=\alpha\beta/(\alpha+\beta-1).$$

The symmetry of this result shows that a similar argument would also give

$$\overline{AP'} = \frac{\alpha\beta}{\alpha+\beta-1}\,\overline{AC},$$

if P' is the point where TS meets AC. Hence P and P' are the same point.]

7. $ABCD$ is a parallelogram, and M the middle point of AB. Show that DM trisects and is trisected by AC*.

28. Multiplication of displacements. So far we have made no attempt to attach any meaning whatever to the notion of the *product* of two displacements. The only kind of multiplication which we have considered is that in which a displacement is multiplied by *a mere number*. The expression

$$[x, y] \times [x', y']$$

so far means nothing, and we are at liberty to define it to mean anything we like. It is, however, fairly clear that if any definition of such a product is to be of any use, the product of two displacements *must itself be a displacement.*

We might, for example, define it as being equal to

$$[x + x', y + y'];$$

in other words, we might agree that the product of two displacements was to be always equal to their sum. But there would be two serious objections to such a definition. In the first place our definition would be futile. We should only be introducing a new method of expressing something which we can perfectly well express without it. In the second place our definition would be inconvenient and misleading for the following reasons. If α is a real number, we have already defined $\alpha[x, y]$ as $[\alpha x, \alpha y]$. Now, as we saw in § 25, the real number α may itself from one point of view be regarded as a displacement, viz. the displacement $[\alpha]$ along the axis OX, or, in our later notation, the displacement $[\alpha, 0]$. It is therefore, if not absolutely necessary, at any rate most desirable, that our definition should be such that

$$[\alpha, 0][x, y] = [\alpha x, \alpha y],$$

and the suggested definition does not give this result.

A more reasonable definition might appear to be

$$[x, y][x', y'] = [xx', yy'].$$

* The two preceding examples are taken from Willard Gibbs' *Vector Analysis*.

But this would give

$$[\alpha, 0]\,[x, y] = [\alpha x, 0],$$

and so this also would be open to the second objection.

In fact, it is by no means obvious what is the best meaning to attach to the product $[x, y]\,[x', y']$. All that is clear is (1) that, if our definition is to be of any use, this product must itself be a displacement whose coordinates depend on x and y, or in other words that we must have

$$[x, y]\,[x', y'] = [X, Y],$$

where X and Y are functions of x, y, x', and y'; (2) that the definition must be such as to agree with the equation

$$[x, 0]\,[x', y'] = [xx', xy'],$$

and (3) that the definition must obey the ordinary commutative, distributive, and associative laws of multiplication, so that

$$[x, y]\,[x', y'] = [x', y']\,[x, y],$$
$$([x, y] + [x', y'])\,[x'', y''] = [x, y]\,[x'', y''] + [x', y']\,[x'', y''],$$
$$[x, y]\,([x', y'] + [x'', y'']) = [x, y]\,[x', y'] + [x, y]\,[x'', y''],$$

and $$[x, y]\,([x', y']\,[x'', y'']) = ([x, y]\,[x', y'])\,[x'', y''].$$

29. The right definition to take is suggested as follows. We know that, if OAB, OCD are two similar triangles, the angles corresponding in the order in which they are written, then

$$OB/OA = OD/OC,$$

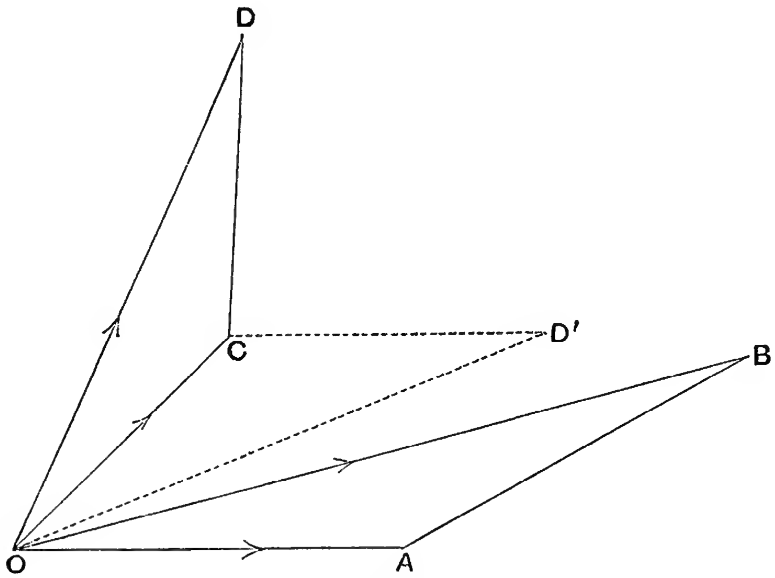


FIG. 28.

or $OB \,.\, OC = OA \,.\, OD$. This suggests that we should try to define multiplication and division of displacements in such a way that

$$\overline{OB}/\overline{OA} = \overline{OD}/\overline{OC}, \quad \overline{OB}\,.\,\overline{OC} = \overline{OA}\,.\,\overline{OD}.$$

Now let

$$\overline{OB} = [x, y], \quad \overline{OC} = [x', y'], \quad \overline{OD} = [X, Y],$$

and suppose that A is the point $(1, 0)$, so that $\overline{OA} = [1, 0]$. Then

$$\overline{OA}\,.\,\overline{OD} = [1, 0]\,[X, Y] = [X, Y],$$

and so

$$[x, y]\,[x', y'] = [X, Y].$$

The product $\overline{OB}\,.\,\overline{OC}$ is therefore to be defined as $\overline{OD}$, D being obtained by constructing on OC a triangle similar to OAB. In order to free this definition from ambiguity, it should be observed that on OC we can describe *two* such triangles, OCD and OCD'. We choose that for which the angle COD is equal to AOB in sign as well as in magnitude. We say that the two triangles are then *similar in the same sense.*

If the polar coordinates of B and C are (ρ, θ) and (σ, ϕ), so that

$$x = \rho \cos\theta, \quad y = \rho \sin\theta, \quad x' = \sigma \cos\phi, \quad y' = \sigma \sin\phi,$$

the polar coordinates of D are evidently $\rho\sigma$ and $\theta + \phi$. Hence

$$X = \rho\sigma \cos(\theta + \phi) = xx' - yy',$$

$$Y = \rho\sigma \sin(\theta + \phi) = xy' + yx'.$$

The required definition is therefore

$$[x, y]\,[x', y'] = [xx' - yy', xy' + yx'] \quad \text{............(6)}.$$

We observe (1) that if $y = 0$, $X = xx'$, $Y = xy'$, as we desired; (2) that the right-hand side is not altered if we interchange x and x', and y and y', so that

$$[x, y]\,[x', y'] = [x', y']\,[x, y];$$

and (3) that

$$\begin{aligned} \{[x, y] + [x', y']\}\,[x'', y''] &= [x + x', y + y']\,[x'', y''] \\ &= [(x + x')\,x'' - (y + y')\,y'', (x + x')\,y'' + (y + y')\,x''] \\ &= [xx'' - yy'', xy'' + yx''] + [x'x'' - y'y'', x'y'' + y'x''] \\ &= [x, y]\,[x'', y''] + [x', y']\,[x'', y'']. \end{aligned}$$

Similarly we can verify that all the equations at the end of § 28 are satisfied. Thus the definition (6) fulfils all the requirements which we made of it in § 28.

Example. Show directly from the geometrical definition given above that multiplication of displacements obeys the commutative and distributive laws. [Take the commutative law for example. The product $\overline{OB}\,.\,\overline{OC}$ is $\overline{OD}$ (Fig. 28), COD being similar to AOB. To construct the product $\overline{OC}\,.\,\overline{OB}$ we should have to *construct on OB a triangle BOD_1 similar to AOC*; and so what we want to prove is that D and D_1 coincide, or that BOD is similar to AOC. This is an easy piece of elementary geometry.]

30. Complex numbers. Just as to a displacement $[x]$ along OX correspond a point (x) and a real number x, so to a displacement $[x, y]$ in the plane correspond a point (x, y) and *a pair of real numbers* x, y.

We shall find it convenient to denote this pair of real numbers x, y by the symbol

$$x + yi.$$

The *reason* for the choice of this notation will appear later. For the present the reader must regard $x + yi$ as *simply another way of writing* $[x, y]$. The expression $x + yi$ is called a *complex number*.

We proceed next to define *equivalence, addition*, and *multiplication* of complex numbers. To every complex number corresponds a displacement. Two complex numbers are equivalent if the corresponding displacements are equivalent. The sum or product of two complex numbers is the complex number which corresponds to the sum or product of the two corresponding displacements. Thus

$$x + yi = x' + y'i \quad \textit{if} \quad x = x',\ y = y' \quad \text{...........(1)},$$

$$(x + yi) + (x' + y'i) = (x + x') + (y + y')\,i \quad \text{.........(2)},$$

$$(x + yi)\,(x' + y'i) = xx' - yy' + (xy' + yx')\,i \quad \text{.........(3)}.$$

In particular, if α is any real number, $\alpha\,(x + yi) = \alpha x + \alpha yi$.

The complex numbers of the particular form $x + 0i$ may be regarded as equivalent to the corresponding real numbers x; thus

$$x + 0i = x,$$

and in particular $0 + 0i = 0$.

Positive integral powers and polynomials of complex numbers are then defined as in ordinary algebra. Thus, by putting $x = x'$, $y = y'$ in (3), we obtain

$$(x + yi)^2 = (x + yi)(x + yi) = x^2 - y^2 + 2xyi,$$

$$(x + yi)^2 + 2(x + yi) + 3 = x^2 - y^2 + 2x + 3 + (2xy + 2y)i.$$

The reader will easily verify for himself that addition and multiplication of complex numbers obey the ordinary laws of algebra, expressed by the equations

$$x + yi + (x' + y'i) = (x' + y'i) + (x + yi),$$

$$\{(x + yi) + (x' + y'i)\} + (x'' + y''i) = (x + yi) + \{(x' + y'i) + (x'' + y''i)\},$$

$$(x + yi)(x' + y'i) = (x' + y'i)(x + yi),$$

$$(x + yi)\{(x' + y'i) + (x'' + y''i)\} = (x + yi)(x' + y'i) + (x + yi)(x'' + y''i),$$

$$\{(x + yi) + (x' + y'i)\}(x'' + y''i) = (x + yi)(x'' + y''i) + (x' + y'i)(x'' + y''i),$$

$$(x + yi)\{(x' + y'i)(x'' + y''i)\} = \{(x + yi)(x' + y'i)\}(x'' + y''i),$$

the proofs of these equations being practically the same as those of the corresponding equations for the corresponding displacements.

Subtraction and division of complex numbers are defined as in ordinary algebra. Thus we may define $(x + yi) - (x' + y'i)$ as

$$(x + yi) + \{-(x' + y'i)\} = x + yi + (-x' - y'i) = (x - x') + (y - y')i;$$

or again, as the number $\xi + \eta i$ such that

$$(x' + y'i) + (\xi + \eta i) = x + yi,$$

which leads to the same result.

And $(x + yi)/(x' + y'i)$ is defined as being the complex number $\xi + i\eta$ such that

$$(x' + y'i)(\xi + \eta i) = x + yi,$$

or

$$x'\xi - y'\eta + (x'\eta + y'\xi)i = x + yi,$$

or

$$x'\xi - y'\eta = x, \quad x'\eta + y'\xi = y \quad \text{..............(4).}$$

Solving these equations for ξ and η, we obtain

$$\xi = \frac{xx' + yy'}{x'^2 + y'^2}, \quad \eta = \frac{yx' - xy'}{x'^2 + y'^2}.$$

This solution fails if x' and y' are both zero, i.e. if $x' + y'i = 0$. Thus subtraction is always possible; division is always possible unless the divisor is zero.

Examples. (1) From a geometrical point of view the problem of the division of the displacement $\overline{OB}$ by $\overline{OC}$ is that of finding D so that the triangles COB, AOD are similar, and this is evidently possible (and the solution unique) *unless C coincides with O*, or $\overline{OC}=0$.

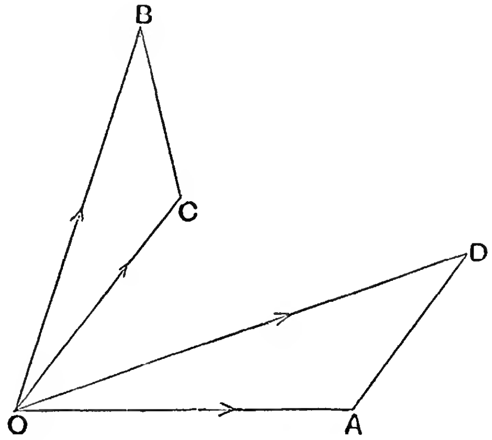


FIG. 29.

(2) The numbers $x+yi$, $x-yi$ are said to be *conjugate*. Verify that

$$(x+yi)(x-yi)=x^2+y^2,$$

so that the product of two conjugate numbers is real, and that

$$\frac{x+yi}{x'+y'i}=\frac{(x+yi)(x'-y'i)}{(x'+y'i)(x'-y'i)}$$
$$=\frac{xx'+yy'+i(x'y-xy')}{x'^2+y'^2}.$$

31. One most important property of real numbers is that known as *the factor theorem*, which asserts that *the product of two numbers cannot be zero unless one of the two is itself zero.* To prove that this is also true of complex numbers we put $x=0$, $y=0$ in the equations (4) of the preceding section. Then

$$x'\xi-y'\eta=0, \quad x'\eta+y'\xi=0.$$

These equations give $\xi=0$, $\eta=0$, i.e.

$$\xi+\eta i=0,$$

unless $x'=0$ and $y'=0$, or $x'+y'i=0$. Thus $x+yi$ cannot vanish unless either $x'+y'i$ or $\xi+\eta i$ vanishes.

32. The equation $i^2=-1$. We agreed to use, instead of $x+0i$, the simpler notation x. Similarly, instead of $0+yi$, we shall use yi. The particular complex number $1i$ we shall denote simply by i. It is the number which corresponds to *a unit displacement along OY.* Also

$$i^2=ii=(0+1i)(0+1i)=(0.0-1.1)+(0.1+1.0)i=-1.$$

Similarly $(-i)^2=-1$. Thus the complex numbers $\pm i$ satisfy the equation $x^2=-1$.

Now the reader will easily satisfy himself that the upshot of the rules for addition and multiplication of complex numbers is this, that *we operate with complex numbers in exactly the same*

way as with real numbers, treating the symbol i as itself a number, but replacing the product $ii = i^2$ by -1 whenever it occurs. Thus, for example,

$$\begin{aligned}(x + yi)(x' + y'i) &= xx' + xy'i + yx'i + yy'i^2 \\ &= (xx' - yy') + (xy' + yx')i.\end{aligned}$$

33. The geometrical interpretation of multiplication by i. Since

$$(x + iy)i = -y + ix,$$

it follows that if $x + iy$ corresponds to $\overline{OP}$, and OQ is drawn equal to OP, and so that POQ is a positive right angle, then $(x + iy)i$ corresponds to $\overline{OQ}$. In other words, *multiplication of a complex number by i turns the corresponding displacement through a right angle.*

We might, had we so chosen, have started from this point of view. We might have regarded x as a length measured along OX, and xi as the same length measured along OY, and regarded i as a symbol of operation equivalent to turning the length x through a right angle round O. We should then naturally have been led to regard $xi^2 = xii$ as denoting the result of *twice* turning x through a right angle. The result of this is to bring it into a position again lying along OX but pointing in the opposite direction, so that we should have been led to the equation

$$xi^2 = -x.$$

Then, denoting $1i$ (a unit length along OY) simply by i, we should have found $i^2 = -1$, and the rule for multiplication of complex numbers would have followed immediately.

34. The equations $x^2 + 1 = 0$, $ax^2 + 2bx + c = 0$. There is no real number x such that $x^2 + 1 = 0$; this is expressed by saying that the equation has *no real roots.* But, as we have just seen, the two complex numbers $\pm i$ satisfy this equation. We express this by saying that the equation has *the two complex roots* $\pm i$. Since i satisfies $x^2 = -1$, it is sometimes written in the form $\sqrt{(-1)}$.

Complex numbers are sometimes called *imaginary,* to distinguish them from *real* numbers. The expression is by no means a happily chosen one, but it is firmly established and

has to be accepted. It cannot, however, be too strongly impressed upon the reader that in reality an 'imaginary number' is neither 'imaginary' nor 'a number' at all. The 'real' numbers would be better described as 'common' or 'ordinary' numbers; they are the numbers of arithmetic. A 'complex' or 'imaginary number' is really not a number at all, but, as should be clear from the following discussion, *a pair of numbers* (x, y), united symbolically, for purposes purely of convenience, in the form $x + yi$. And such a pair of numbers is no less 'real' than any ordinary number such as $\frac{1}{2}$, or than the paper on which this is printed, or than the Solar System.

In reality

$$i = 0 + 1i$$

stands for the pair of numbers $(0, 1)$, and may be represented geometrically by a point or by the displacement $[0, 1]$. And when we say that i is a root of the equation $x^2 + 1 = 0$, what we mean is simply that we have defined a method of combining such pairs of numbers (or displacements) which we call 'multiplication,' and which, when we so combine $(0, 1)$ with itself, gives the result $(-1, 0)$.

Now let us consider the more general equation

$$ax^2 + 2bx + c = 0,$$

where a, b, c are real numbers.

If $b^2 > ac$, the ordinary method of solution gives two real roots

$$\{-b \pm \sqrt{(b^2 - ac)}\}/a.$$

If $b^2 < ac$, the equation has no real roots. It may be written in the form

$$\{x + (b/a)\}^2 = -(ac - b^2)/a^2,$$

an equation which is evidently satisfied if we substitute for $x + (b/a)$ the complex number $\pm i\sqrt{(ac - b^2)}/a$. We express this by saying that the equation has *the two complex roots*

$$\{-b \pm i\sqrt{(ac - b^2)}\}/a.$$

If we agree as a matter of convention to say that when $b^2 = ac$ (in which case the equation is satisfied by *one* value of x only, viz. $-b/a$), the equation has *two equal roots*, we can say that *a quadratic equation with real coefficients has two roots in all*

cases, two distinct real roots, two equal real roots, or two distinct complex roots.

The question is naturally suggested whether a quadratic equation may not, when complex roots are once admitted, have more than two roots. It is easy to see that this is not possible. In fact, its impossibility may be proved by precisely the same chain of reasoning as is used in elementary algebra to prove that an equation of the nth degree cannot have more than n real roots. Let $z = x + yi$, and let $f(z)$ denote any polynomial in z*, with real or complex coefficients. Then we prove in succession:

(1) that the remainder, when $f(z)$ is divided by $z - a$ (a being any real or complex number), is $f(a)$;

(2) if a is a root of the equation $f(z) = 0$, then $f(z)$ is divisible by $z - a$;

(3) if $f(z)$ is of the nth degree, and $f(z) = 0$ has the n roots $a_1, a_2, \dots, a_n$, then

$$f(z) \equiv A(z - a_1)(z - a_2) \dots (z - a_n),$$

where A is a constant (real or complex), in fact the coefficient of z^n in $f(z)$. From the last result it follows at once that $f(z)$ cannot have more than n roots.

We conclude that a quadratic equation with real coefficients has exactly two roots. We shall see later on that a similar theorem is true for an equation of any degree and with either real or complex coefficients: *an equation of the nth degree has exactly n roots.* The only point in the proof which presents any difficulty is the first, viz. the proof that any equation must have *at least one* root. This we must postpone for the present. We may, however, at once call attention to one very interesting result of this theorem. In the theory of number we start from the positive integers, and from the ideas of addition and multiplication, and the converse operations of subtraction and division. We find that these operations are not always possible unless we admit new kinds of numbers. We can only attach a meaning to $3 - 7$ if we admit *negative* numbers, or to $\frac{3}{7}$ if we admit *rational fractions.* When we extend our list of arithmetical operations so as to include root extraction and the solution of equations, we find

* A polynomial in $z = x + yi$ is of course defined in exactly the same way as a polynomial in x, i.e. as an expression of the form $a_0 z^n + a_1 z^{n-1} + \dots + a_n$.

that some of them, such as the extraction of the square root of a number which (like 2) is not a perfect square, are not possible unless we widen our conception of a number, and admit the irrational numbers of Chap. I.

Others, such as the extraction of the square root of -1, are not possible unless we go still further, and admit the *complex* numbers of this chapter. And it would not be unnatural to suppose that, when we come to consider equations of higher degree, some might prove to be insoluble even by the aid of complex numbers, and that thus we might be led to the considerations of higher and higher types of, so to say, *hyper-complex* numbers. The fact that any algebraical equation whatever can be solved by means of ordinary complex numbers shows that this is not the case. The application of any of the ordinary algebraical operations to complex numbers will yield only complex numbers. In technical language 'the field of the complex numbers is closed for algebraical operations.'

Before we pass on to other matters, let us add that all theorems of elementary algebra which are proved merely by the application of the rules of addition and multiplication are true *whether the numbers which occur in them are real or complex,* since the rules referred to apply to complex as well as real numbers. For example, we know that if α and β are the roots of

$$ax^2 + 2bx + c = 0,$$

then

$$\alpha + \beta = -(2b/a), \quad \alpha\beta = (c/a).$$

Similarly, if α, β, γ are the roots of

$$ax^3 + 3bx^2 + 3cx + d = 0,$$

then

$$\alpha + \beta + \gamma = -(3b/a), \quad \beta\gamma + \gamma\alpha + \alpha\beta = (3c/a), \quad \alpha\beta\gamma = -(d/a).$$

All such theorems as these are true whether $a, b, \dots \alpha, \beta, \dots$ are real or complex.

35. The Argand diagram. Let P be the point (x, y) in Fig. 30, r, θ its polar coordinates, so that

$$x = r\cos\theta, \quad y = r\sin\theta, \quad r = \sqrt{(x^2+y^2)}, \quad \cos\theta : \sin\theta : 1 :: x : y : r.$$

We shall denote the complex number $x + yi$ by z, and we

shall call z the *complex variable.* We shall call P *the point* z, or the point corresponding to z; and z the *argument of* P. We shall call x the *real part,* y the *imaginary part,* r the *modulus,* and θ the *amplitude* of z, and we shall write

$$x = R(z), \quad y = I(z),$$
$$r = |z|, \quad \theta = \text{am}\, z.$$

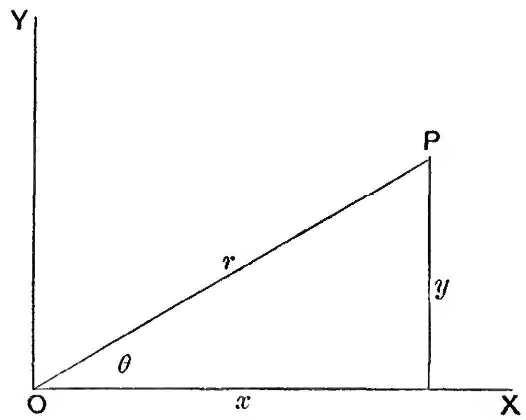


Fig. 30.

It should be observed that r is essentially *positive* (except when $z = 0$).

When $y = 0$ we shall say that z *is real,* when $x = 0$ that z *is purely imaginary.* Two numbers $x + yi$, $x - yi$ which differ only in the signs of their imaginary parts, we shall call *conjugate.* It will be observed that the sum $(2x)$ of two conjugate numbers and their product $(x^2 + y^2)$ are both real, that their moduli $(\sqrt{x^2 + y^2})$ are equal, and that their product is equal to the square of the modulus of either. The roots of a quadratic with real coefficients, for example, are conjugate, when not real.

It must be observed that θ or am z is a many-valued function of z, having an infinity of values differing by multiples of 2π. Any one of its values is an angle by turning through which about O a line originally lying along OX will come to lie along OP. We shall denote that one of these values which lies between $-\pi$ and $+\pi$ as the *principal value* of the amplitude of z. This definition is unambiguous except when one of the values is π, in which case $-\pi$ is also a value. In this case we must make some special provision as to which value is to be regarded as the principal value. In general, when we speak of the amplitude of z we shall, unless the contrary is stated, mean the principal value of the amplitude.

Complex numbers were first studied from a geometrical point of view by Wessel, Gauss and Argand, and the figure is usually known as the Argand diagram.

36. De Moivre's Theorem. The following statements follow immediately from the definitions of addition and multiplication.

(1) The real (or imaginary) part of the sum of two complex numbers is the sum of their real (or imaginary) parts.

(2) The modulus of the product of two complex numbers is the product of their moduli.

(3) One value of the amplitude of the product of two complex numbers is the sum of their amplitudes.

It should be observed that it is not true that the principal value of $\mathrm{am}\,(zz')$ is the sum of the principal values of $\mathrm{am}\, z$ and $\mathrm{am}\, z'$. For example, if $z=z'=-1+i$, the principal values of the amplitudes of z and z' are each $\frac{3}{4}\pi$. But $zz'=-2i$, and the principal value of $\mathrm{am}\,(zz')$ is $-\frac{1}{2}\pi$ and not $\frac{3}{2}\pi$.

The two last theorems may be expressed in the equation

$$r(\cos\theta + i\sin\theta) \times \rho(\cos\phi + i\sin\phi) \\ = r\rho\{\cos(\theta+\phi) + i\sin(\theta+\phi)\},$$

which may be proved at once by multiplying out and using the ordinary trigonometrical formulae for $\cos(\theta+\phi)$ and $\sin(\theta+\phi)$. More generally

$$r_1(\cos\theta_1 + i\sin\theta_1) \times r_2(\cos\theta_2 + i\sin\theta_2) \times \dots \times r_n(\cos\theta_n + i\sin\theta_n) \\ = r_1 r_2 \dots r_n\{\cos(\theta_1+\theta_2+\dots+\theta_n) + i\sin(\theta_1+\theta_2+\dots+\theta_n)\}.$$

A particularly interesting case is that in which

$$r_1 = r_2 = \dots = r_n = 1, \quad \theta_1 = \theta_2 = \dots = \theta_n = \theta.$$

We then obtain the equation

$$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta,$$

where n is any positive integer: a result known as *De Moivre's Theorem**.

Again, if
$$z = r(\cos\theta + i\sin\theta), \\ 1/z = (\cos\theta - i\sin\theta)/r.$$

Thus the modulus of the reciprocal of z is the reciprocal of the modulus of z, and the amplitude of the reciprocal is the amplitude of z with its sign changed. Hence we deduce from (2) and (3):

(4) The modulus of the quotient of two complex numbers is the quotient of their moduli.

(5) One value of the amplitude of the quotient of two complex numbers is the difference of their amplitudes.

* It will sometimes be convenient, for the sake of brevity, to denote $\cos\theta + i\sin\theta$ by $\mathrm{Cis}\,\theta$: in this notation, suggested by Profs. Harkness and Morley, De Moivre's theorem is expressed by the equation $(\mathrm{Cis}\,\theta)^n = \mathrm{Cis}\, n\theta$.

Again $$(\cos\theta + i\sin\theta)^{-n} = (\cos\theta - i\sin\theta)^n$$
$$= \{\cos(-\theta) + i\sin(-\theta)\}^n$$
$$= \cos(-n\theta) + i\sin(-n\theta).$$

Hence *De Moivre's Theorem holds for all integral values of n, positive or negative.* A large number of important applications of this theorem will be given later on in this chapter (§§ 38 *et seq.*).

To the theorems (1)—(5) we may add the following theorem, which is also of very great importance.

(6) The modulus of the sum of any number of complex quantities is *not greater than* the sum of their moduli.

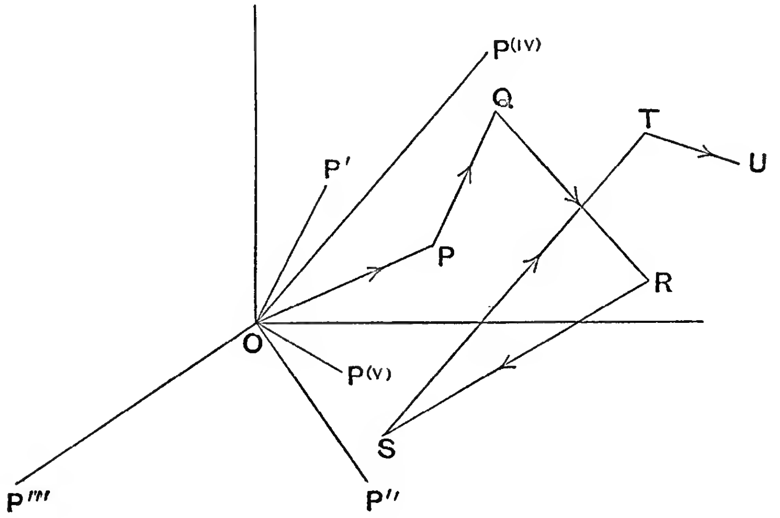


FIG. 31.

Let $\overline{OP}$, $\overline{OP'}$, ... be the displacements corresponding to the various complex quantities. Draw PQ equal and parallel to OP', QR equal and parallel to OP'', and so on. Finally we reach a point U, such that

$$\overline{OU} = \overline{OP} + \overline{OP'} + \overline{OP''} + \dots.$$

The length OU is the modulus of the sum of the complex quantities, whereas the sum of their moduli is the total length of the broken line $OPQR\dots U$. The truth of the theorem is now obvious (see also Ex. XXIII. 1).

37. We add some theorems concerning rational functions of complex numbers. A *rational function* of the complex variable z is defined exactly as is a rational function of a real variable x, viz. as the quotient of two polynomials in z.

THEOREM 1. *If $R(x+yi)$ is a rational function of $x+yi$, it can be reduced to the form $X+Yi$, where X and Y are rational functions of x and y with real coefficients.*

In the first place it is evident that any polynomial $P(x+yi)$ can be reduced, in virtue of the definitions of addition and multiplication, to the form $A+Bi$, where A and B are polynomials in x and y with real coefficients. Similarly $Q(x+yi)$ can be reduced to the form $C+Di$. Hence

$$R(x+yi)=P(x+yi)/Q(x+yi)$$

can be expressed in the form

$$\begin{aligned}(A+Bi)/(C+Di) &= (A+Bi)(C-Di)/(C+Di)(C-Di)\\ &= \frac{AC+BD}{C^2+D^2}+\frac{BC-AD}{C^2+D^2}i,\end{aligned}$$

which proves the theorem.

THEOREM 2. *If $R(x+yi)=X+Yi$, R denoting a rational function as before, but with* **real** *coefficients, then $R(x-yi)=X-Yi$.*

In the first place this is easily verified for a power $(x+yi)^n$ by actual expansion.

It follows by addition that the theorem is true for any polynomial with real coefficients. Hence, in the notation used above,

$$R(x-yi)=\frac{A-Bi}{C-Di}=\frac{AC+BD}{C^2+D^2}-\frac{BC-AD}{C^2+D^2}i,$$

the reduction being the same as before except that the sign of i is changed throughout. It is evident that results similar to those of Theorems 1 and 2 hold for functions of any number of complex variables.

THEOREM 3. *The roots of an equation*

$$a_0z^n+a_1z^{n-1}+\dots+a_n=0,$$

whose coefficients are real, may, in so far as they are not themselves real, be arranged in conjugate pairs.

For it follows from Theorem 2 that if $x+yi$ is a root, so is $x-yi$. A particular case of this theorem is the result (§ 34) that the roots of a quadratic equation with real coefficients are either real or conjugate.

This theorem is sometimes stated as follows—*in an equation*

with real coefficients complex roots occur in conjugate pairs. It should be compared with the result of Exs. IV. 10, which may be stated as follows—*in an equation with rational coefficients irrational roots occur in conjugate pairs**.

Examples XXIII. 1. Prove theorem (6) of § 36 directly from the definitions and without the aid of geometrical considerations.

[First, to prove that $|z+z'| \leqq |z|+|z'|$ is to prove that

$$\cancel{(x+y)}^2+\cancel{(x'+y')}^2 \leqq \{\sqrt{(x^2+y^2)}+\sqrt{(x'^2+y'^2)}\}^2.$$

The theorem is then easily extended to the general case.]

2. The one and only case in which

$$|z|+|z'|+\ldots=|z+z'+\ldots|,$$

is that in which the numbers $z, z', \ldots$ have all the same amplitude. Prove this both geometrically and analytically.

3. The modulus of the sum of any number of complex numbers is not less than the sum of their real (or imaginary) parts.

4. If the sum and product of two complex numbers are both real the two numbers must either be real or conjugate.

5. If $\qquad a+b\sqrt{2}+(c+d\sqrt{2})i=A+B\sqrt{2}+(C+D\sqrt{2})i,$

where a, b, c, d, A, B, C, D are real rational numbers, then

$$a=A, \quad b=B, \quad c=C, \quad d=D.$$

6. Express the following numbers in the form $A+Bi$, where A and B are real numbers: $(1+i)^2$, $(1-i)^2$, $(3-2i)/(2+3i)$, $(\lambda+\mu i)/(\lambda-\mu i)$,

$$\left(\frac{1+i}{1-i}\right)^2, \quad \left(\frac{1-i}{1+i}\right)^2, \quad \left(\frac{1+i}{1-i}\right)^3-\left(\frac{1-i}{1+i}\right)^3, \quad \left(\frac{\lambda+\mu i}{\lambda-\mu i}\right)^2-\left(\frac{\lambda-\mu i}{\lambda+\mu i}\right)^2,$$

λ and μ denoting any real numbers.

7. Express the following functions of $z=x+yi$ in the form $X+Yi$, where X and Y are real functions of x and y: z^2, z^3, z^n, $1/z$, $z+(1/z)$, $(1+z)/(1-z)$, $(\alpha+\beta z)/(\gamma+\delta z)$, $\alpha, \beta, \gamma, \delta$ denoting real numbers.

8. Find the moduli of the numbers and functions in the two preceding examples.

9. The two lines joining the points $z=a$, $z=b$ and $z=c$, $z=d$ will be perpendicular if

$$\operatorname{am}\left(\frac{a-b}{c-d}\right)=\pm\tfrac{1}{2}\pi;$$

i.e. if $(a-b)/(c-d)$ is purely imaginary. What is the condition that the lines should be parallel?

10. The three angular points of a triangle are given by $z=\alpha$, $z=\beta$, $z=\gamma$, where α, β, γ are complex quantities. Establish the following propositions:

(i) *The centre of gravity is given by* $z=\frac{1}{3}(\alpha+\beta+\gamma)$.

(ii) *The circum-centre is given by* $|z-\alpha|=|z-\beta|=|z-\gamma|$.

* The numbers $a+\sqrt{b}$, $a-\sqrt{b}$, where a, b are rational, are sometimes said to be 'conjugate.'

(iii) *The three perpendiculars from the angular points on the opposite sides meet in a point given by*

$$R\left(\frac{z-a}{\beta-\gamma}\right)=R\left(\frac{z-\beta}{\gamma-a}\right)=R\left(\frac{z-\gamma}{a-\beta}\right)=0.$$

[If A, B, C are the vertices, and P any point z, the condition that AP should be perpendicular to BC (Ex. 9) is that $(z-a)/(\beta-\gamma)$ should be purely imaginary, or that $R(z-a)\,R(\beta-\gamma)+I(z-a)\,I(\beta-\gamma)=0$.

This equation and the two similar equations, obtained by permuting a, β, γ cyclically, are satisfied by the same value of z, as appears from the fact that the sum of the three left-hand sides is zero (so that the third equation is a consequence of the first two). This proves the theorem.]

(iv) *There is a point P inside the triangle such that*

$$CBP=ACP=BAP=\omega.$$

Also $$\cot\omega=\cot A+\cot B+\cot C.$$

[From the equations*

$$\omega=CBP=\operatorname{am}(z-\beta)-\operatorname{am}(\gamma-\beta),$$
$$\cot\operatorname{am}(z-\beta)=\{R(z-\beta)/I(z-\beta)\},\ \text{etc.},$$

we deduce

$$\cot\omega\,\{I(z-\beta)\,R(\gamma-\beta)-R(z-\beta)\,I(\gamma-\beta)\}$$
$$=R(z-\beta)\,R(\gamma-\beta)+I(z-\beta)\,I(\gamma-\beta)\quad\ldots\ldots\ldots(1).$$

This equation, and the two similar equations obtained by permuting a, β, γ cyclically, suffice to determine $\cot\omega$ and the real and imaginary parts of z. If we add the three equations z disappears and we are left with

$$\cot\omega\,\Sigma\,\{I(\beta)\,R(\gamma)-R(\beta)\,I(\gamma)\}$$
$$=\Sigma\,\{R(\beta)\,R(\gamma-\beta)+I(\beta)\,I(\gamma-\beta)\}\quad\ldots\ldots\ldots\ldots(2),$$

the sign of summation referring to the three terms produced by cyclical interchange of a, β, γ.

Now $$\cot A=\cot\{\operatorname{am}(\gamma-a)-\operatorname{am}(\beta-a)\}$$
$$=\frac{R(\gamma-a)\,R(\beta-a)+I(\gamma-a)\,I(\beta-a)}{I(\gamma-a)\,R(\beta-a)-R(\gamma-a)\,I(\beta-a)},$$

and a little reduction shows that the denominator of this fraction is equal to the coefficient of $\cot\omega$ in equation (2), with its sign changed. Hence we can deduce that $\cot\omega=\cot A+\cot B+\cot C$.]

11. The two triangles whose vertices are the points a, b, c and x, y, z respectively will be similar if

$$\begin{vmatrix}1 & 1 & 1\\ a & b & c\\ x & y & z\end{vmatrix}=0.$$

[The condition required is that $\overline{AB}/\overline{AC}=\overline{XY}/\overline{XZ}$ (large letters denoting the points whose arguments are the corresponding small letters), or $(b-a)/(c-a)=(y-x)/(z-x)$, which is the same as the given condition.]

* We suppose that as we go round the triangle in the direction ABC we leave it on our left.

12. Deduce from the last example that if the points x, y, z are collinear we can find *real* numbers α, β, γ such that $\alpha+\beta+\gamma=0$ and $\alpha x+\beta y+\gamma z=0$, and conversely (cf. Exs. XXII. 4). [Use the fact that in this case the triangle formed by x, y, z is similar to a certain line-triangle on the axis OX, and apply the result of the last example.]

13. **The general equation of the first degree with complex coefficients.** The equation $az+b=0$ has the one solution $z=-(b/a)$, unless $a=0$. If we put

$$a=\alpha+i\alpha', \quad b=\beta+i\beta', \quad z=x+iy,$$

and equate real and imaginary parts, we obtain two equations to determine the two real quantities x and y. The equation will have a *real* root if $y=0$, which gives $\alpha x+\beta=0$, $\alpha' x+\beta'=0$, and the condition that these equations should be consistent is $\alpha\beta'-\alpha'\beta=0$.

14. **The general quadratic equation with complex coefficients.** This equation is

$$(a+iA)z^2+2(b+iB)z+(c+iC)=0.$$

Unless a and A are both zero we can divide through by $a+iA$. Hence we may consider

$$z^2+2(b+iB)z+(c+iC)=0 \quad \text{........................(1)},$$

as the standard form of our equation. Putting $z=x+iy$ and equating real and imaginary parts we obtain a pair of simultaneous equations for x and y, viz.

$$x^2-y^2+2(bx-By)+c=0, \quad 2xy+2(by+Bx)+C=0.$$

If we put

$$x+b=\xi, \quad y+B=\eta, \quad b^2-B^2-c=h, \quad 2bB-C=k,$$

these equations become $\quad \xi^2-\eta^2=h, \quad 2\xi\eta=k.$

Squaring and adding we obtain

$$\xi^2+\eta^2=\sqrt{(h^2+k^2)}, \quad \xi=\pm\sqrt{\tfrac{1}{2}\{\sqrt{(h^2+k^2)}+h\}}, \quad \eta=\pm\sqrt{\tfrac{1}{2}\{\sqrt{(h^2+k^2)}-h\}}.$$

We must choose the signs so that $\xi\eta$ has the sign of k: i.e. if k is positive we must take like signs, if k is negative unlike signs.

Conditions for equal roots. The two roots can only be equal if both the square roots above vanish, i.e. if $h=0$, $k=0$, or if $c=b^2-B^2$, $C=2bB$. These conditions are equivalent to the single condition $c+iC=(b+iB)^2$ which obviously expresses the fact that the left-hand side of (1) is a perfect square.

Condition for a real root. If $x^2+2(b+iB)x+(c+iC)=0$, where x is real, then $x^2+2bx+c=0$, $2Bx+C=0$. Eliminating x we find that the required condition is $\quad C^2-4bBC+4cB^2=0.$

Condition for a purely imaginary root. This is easily found to be

$$C^2-4bBC-4b^2c=0.$$

Conditions for a pair of conjugate complex roots. Since the sum and the product of two conjugate complex quantities are both real, $b+iB$ and $c+iC$ must both be real, i.e. $B=0$, $C=0$. Thus the equation (1) can have a pair of conjugate complex roots only if its coefficients are real. The reader should

verify this conclusion by means of the explicit expressions of the roots. Moreover, even in this case, if $b^2 \geqq c$ the roots will be real. Hence for a pair of conjugate roots we must have $B=0$, $C=0$, $b^2<c$.

15. **The Cubic equation.** Consider the cubic equation

$$z^3+3Hz+G=0,$$

where G and H are complex quantities, it being given that the equation has (a) a real root, (b) a purely imaginary root, (c) a pair of conjugate roots.

If $H=\lambda+i\mu$, $G=\rho+i\sigma$, we arrive at the following conclusions.

(a) *A real root.* If $\mu \neq 0$ the real root is $-\sigma/3\mu$, and $\sigma^3+27\lambda\mu^2\sigma-27\mu^3\rho=0$. On the other hand, if $\mu=0$, we must also have $\sigma=0$, and the coefficients of the equation are real. In this case there may be three real roots.

(b) *A purely imaginary root.* If $\mu \neq 0$ the purely imaginary root is $(\rho/3\mu)i$, and $\rho^3-27\lambda\mu^2\rho-27\mu^3\sigma=0$. If $\mu=0$, then also $\rho=0$, and the root is iy, where y is given by the equation $y^3-3\lambda y-\sigma=0$, which has real coefficients. In this case there may be three purely imaginary roots.

(c) *A pair of conjugate roots.* Let these be $x \pm yi$. Then since the sum of the three roots is zero the third root must be $-2x$. From the relations between the coefficients and the roots of an equation we deduce

$$y^2-3x^2=3H, \quad 2x(x^2+y^2)=G.$$

Hence G and H must both be real.

In each case we can either find a root (in which case the equation can be reduced to a quadratic by dividing by a known factor) or we can reduce the solution of the equation to the solution of a cubic equation with real coefficients.

16. The cubic equation $x^3+a_1x^2+a_2x+a_3=0$, where $a_1=A_1+iA_1'$, ... has a pair of conjugate imaginary roots. Prove that provided $A_3' \neq 0$ the remaining root is $-A_1'a_3/A_3'$, and two identical relations hold between $A_1, A_1', A_2, \ldots$. Examine the case in which $A_3'=0$.

17. Prove that if $z^3+3Hz+G=0$ has two imaginary roots, the equation

$$8a^3+6aH-G=0$$

has one real root which is the real part a of the imaginary roots of the original equation; and show that a has the same sign as G.

18. An equation of any order with complex coefficients will *in general* have no real roots, nor pairs of conjugate complex roots. How many conditions must be satisfied by the coefficients in order that the equation should have (a) a real root, (b) a pair of conjugate roots?

19. **Coaxal circles.** In Fig. 32, let a, b, z be the arguments of A, B, P. Then

$$\operatorname{am}\frac{z-b}{z-a}=APB,$$

the principal value of the amplitude being taken, and APB being a positive angle less than π. If the two circles shown in the figure are equal, and

z', z_1, z_1' are the arguments of P', P_1, P_1', and $APB=\theta$, it is easy to see that

$$\text{am}\frac{z'-b}{z'-a}=\pi-\theta,\quad \text{am}\frac{z_1-b}{z_1-a}=-\theta,$$

and
$$\text{am}\frac{z_1'-b}{z_1'-a}=-\pi+\theta.$$

The locus defined by the equation

$$\text{am}\frac{z-b}{z-a}=\theta,$$

where θ is constant, is the arc APB. By writing $\pi-\theta$, $-\theta$, $-\pi+\theta$ for θ we obtain the other three arcs shown.

The system of equations obtained by supposing that θ is a parameter, varying from $-\pi$ to $+\pi$, represents *the system of circles which can be drawn through the points A, B.* It should however be observed that each circle has to be divided into two parts to which correspond different values of θ.

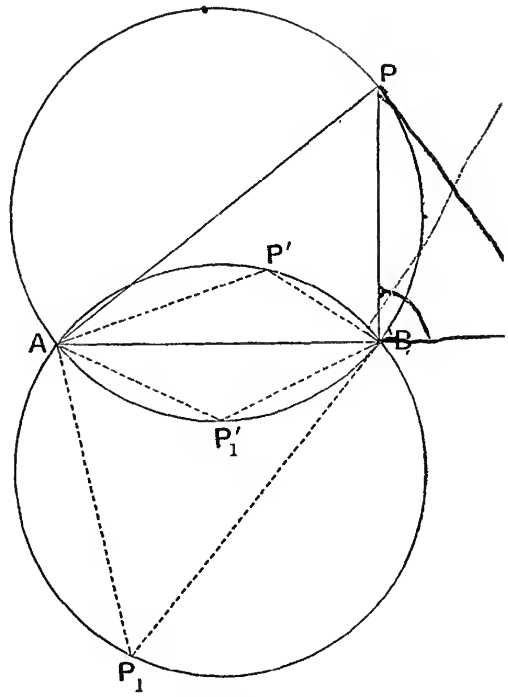


FIG. 32.

20. Now let us consider the equation

$$|(z-b)/(z-a)|=\lambda \quad \ldots\ldots\ldots\ldots\ldots\ldots(1),$$

where λ is a constant.

Take the point K on BA produced so that $KPA=KBP$. Then the triangles KPA, KBP are similar, and so

$$AP/BP=KP/KB=KA/KP=\lambda.$$

Hence $KA/KB=\lambda^2$, and therefore K is a fixed point for all positions of P which satisfy the equation (1). Also $KP^2=KA\,.\,KB=\text{const.}$ Hence *the locus of P is a circle whose centre is K.* For different values of λ the equation (1) therefore represents a system of circles.

Every circle of this system cuts at right angles every circle of the system of Ex. 19. For the equation $KP^2=KA\,.\,KB$ shows that KP is a tangent to the circle APB.

The system of Ex. 19 is *a system of coaxal circles of the common point kind.* The system of Ex. 20 is called *a system of coaxal circles of the limiting point kind*; if λ is very small the circle is a very small circle containing B in its interior, if λ is very large a very small circle containing A in its interior: it is from this fact that the name is derived.

21. **Bilinear Transformations.** Consider the equation

$$z=Z+a\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1),$$

where $z=x+iy$ and $Z=X+iY$ are two complex variables which we may suppose to be represented in two planes xoy, XOY. To every value of z

corresponds one of Z, and conversely. If $a=\alpha+i\beta$, then

$$x=X+\alpha, \quad y=Y+\beta,$$

and to the point (x, y) corresponds the point (X, Y). If (x, y) describes a curve of any kind in its plane, (X, Y) describes a curve in its plane. Thus to any figure in one plane corresponds a figure in the other. A passage of this kind from a figure in the plane xoy to a figure in the plane XOY by means of a relation such as (1) between z and Z is called a *transformation*. In this particular case the relation between corresponding figures is very easily defined. The (X, Y) figure is the same in size, shape, and orientation as the (x, y) figure, but is shifted a distance α to the left, and a distance β downwards. Such a transformation is called a *translation*.

Now consider the equation

$$z=\rho Z \quad \text{.....................................(2)},$$

where ρ is real. This gives $x=\rho X$, $y=\rho Y$. The two figures are similar and similarly situated about their respective origins, but the scale of the (X, Y) figure is $(1/\rho)$ times that of the (x, y) figure. Such a transformation is called a *magnification*.

Finally consider the equation

$$z=(\cos\phi+i\sin\phi)\,Z \quad \text{.............................(3)}.$$

It is clear that $|z|=|Z|$, $\operatorname{am} z=\operatorname{am} Z+\phi$, and that the two figures differ only in that the (X, Y) figure is the (x, y) figure turned about the origin through an angle ϕ in the negative direction. Such a transformation is called a *rotation*.

The general linear transformation

$$z=aZ+b \quad \text{...................................(4)}$$

is a combination of the three transformations (1), (2), (3). For if $|a|=\rho$ and $\operatorname{am} a=\phi$ we can replace (4) by the three equations

$$z=z'+b, \quad z'=\rho Z', \quad Z'=(\cos\phi+i\sin\phi)\,Z.$$

Thus *the general linear transformation is equivalent to the combination of a translation, a magnification, and a rotation.*

Next let us consider the transformation

$$z=1/Z \quad \text{....................................(5)}.$$

If $|Z|=R$ and $\operatorname{am} Z=\Theta$, then $|z|=1/R$ and $\operatorname{am} z=-\Theta$, and to pass from the (x, y) figure to the (X, Y) figure we *invert* the former with respect to o, with unit radius of inversion, and then construct the *image* of the new figure in the axis ox (i.e. the symmetrical figure on the other side of ox). We thus obtain a figure in the (x, y) plane, similar in every respect to the (X, Y) figure.

Finally consider the transformation

$$z=(aZ+b)/(cZ+d) \quad \text{.............................(6)}.$$

This is equivalent to the combination of the transformations

$$z=(a/c)+(bc-ad)(z'/c), \quad z'=1/Z', \quad Z'=cZ+d,$$

i.e. to a certain combination of transformations of the types already considered.

The transformation $$z=(aZ+b)/(cZ+d)$$
is called the *general bilinear transformation.* Solving for Z we obtain
$$Z=(dz-b)/(cz-a).$$

This is the most general type of transformation for which one and only one value of z corresponds to each value of Z, and conversely.

22. *The general bilinear transformation transforms circles into circles.* This may be proved in a variety of ways. We may assume the well known theorem in pure geometry, that *inversion* transforms circles into circles (which may of course in particular cases be *straight lines*). Or we may take the equation
$$x+iy=\frac{(\alpha+i\alpha')(X+iY)+(\beta+i\beta')}{(\gamma+i\gamma')(X+iY)+(\delta+i\delta')},$$
assume that x and y satisfy the equation of a circle, calculate x and y in terms of X and Y, and so find the relation between X and Y by straightforward algebra. Or finally we may use the results of Exs. 19 and 20. This is the best and simplest method. If, e.g., the (x, y) circle is
$$|(z-\sigma)/(z-\rho)|=\lambda,$$
and we substitute for z in terms of Z, we obtain
$$|(Z-\sigma')/(Z-\rho')|=\lambda',$$
where $$\sigma'=-\frac{b-\sigma d}{a-\sigma c},\quad \rho'=-\frac{b-\rho d}{a-\rho c},\quad \lambda'=\lambda\left|\frac{a-\rho c}{a-\sigma c}\right|.$$

23. Consider the transformations $z=1/Z$, $z=(1+Z)/(1-Z)$, and draw the (X, Y) curves which correspond to (1) circles whose centre is the origin, (2) straight lines through the origin, in the (x, y) plane.

24. The condition that the transformation $z=(aZ+b)/(cZ+d)$ should make the circle $x^2+y^2=1$ correspond to a straight line in the (X, Y) plane, is $|a|=|c|$.

25. **Cross ratios.** The cross ratio (z_1z_2, z_3z_4) is defined to be
$$\frac{(z_1-z_3)(z_2-z_4)}{(z_1-z_4)(z_2-z_3)}.$$

If the four points z_1, z_2, z_3, z_4 are on the same line, this agrees with the definition adopted in elementary geometry. There are 24 cross ratios which can be formed from z_1, z_2, z_3, z_4 by permuting the suffixes. These consist of six groups of four equal cross ratios. If one ratio is λ the six distinct cross ratios are λ, $1-\lambda$, $1/\lambda$, $1/(1-\lambda)$, $1-(1/\lambda)$, $\lambda/(\lambda-1)$. The four points are said to be *harmonic* or *harmonically related* if any one of these is equal to -1. In this case the six ratios are -1, -1, $\frac{1}{2}$, $\frac{1}{2}$, 2, 2.

If any cross ratio is real all are real and the four points lie on a circle. For in this case
$$\operatorname{am}\frac{(z_1-z_3)(z_2-z_4)}{(z_1-z_4)(z_2-z_3)}=0\ (\text{or } \pi),$$

so that am $\{(z_1-z_3)/(z_1-z_4)\}$ and am $\{(z_2-z_3)/(z_2-z_4)\}$ are either equal or differ by π (cf. Ex. 19).

If $(z_1z_2,\ z_3z_4)=-1$, we have the two equations

$$\operatorname{am}\frac{z_1-z_3}{z_1-z_4}=\pi+\operatorname{am}\frac{z_2-z_3}{z_2-z_4},\qquad \left|\frac{z_1-z_3}{z_1-z_4}\right|=\left|\frac{z_2-z_3}{z_2-z_4}\right|.$$

The four points A_1, A_2, A_3, A_4 (Fig. 33) lie on a circle, A_1 and A_2 being separated by A_3 and A_4. Also $A_1A_3/A_1A_4=A_2A_3/A_2A_4$. Let O be the middle point of A_3A_4. The equation

$$\frac{(z_1-z_3)(z_2-z_4)}{(z_1-z_4)(z_2-z_3)}=-1$$

may be put in the form

$$(z_1+z_2)(z_3+z_4)=2(z_1z_2+z_3z_4),$$

or, what is the same thing,

$$\{z_1-\tfrac{1}{2}(z_3+z_4)\}\{z_2-\tfrac{1}{2}(z_3+z_4)\}=\{\tfrac{1}{2}(z_3-z_4)\}^2.$$

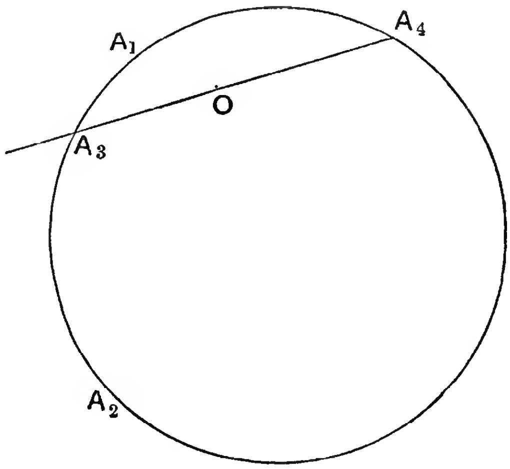


Fig. 33.

But this is equivalent to $\overline{OA_1}\,.\,\overline{OA_2}=\overline{OA_3}^2=\overline{OA_4}^2$. Hence OA_1 and OA_2 make equal angles with A_3A_4, and $OA_1\,.\,OA_2=OA_3^2=OA_4^2$. It will be observed that the relation between the pairs A_1, A_2 and A_3, A_4 is symmetrical. Hence if O' is the middle point of A_1A_2, $O'A_3$ and $O'A_4$ are equally inclined to A_1A_2, and $O'A_3\,.\,O'A_4=O'A_1^2=O'A_2^2$.

26. If the points A_1, A_2 are given by $az^2+2bz+c=0$, and the points A_3, A_4 by $a'z^2+2b'z+c'=0$, and O is the middle point of A_3A_4, and $ac'+a'c-2bb'=0$, then OA_1, OA_2 are equally inclined to A_3A_4 and $OA_1\,.\,OA_2=OA_3^2=OA_4^2$. (*Math. Trip.* 1901.)

[The pairs A_1, A_2 and A_3, A_4 are harmonically related.]

27. **The condition that four points should lie on a circle.** A sufficient condition is that one (and therefore all) of the cross ratios should be real (Ex. 25); this condition is also necessary. Another form of the condition is that it should be possible to choose real quantities α, β, γ such that

$$\begin{vmatrix} 1 & 1 & 1 \\ \alpha & \beta & \gamma \\ z_1z_4+z_2z_3 & z_2z_4+z_3z_1 & z_3z_4+z_1z_2 \end{vmatrix}=0.$$

To prove this we observe that the transformation $Z=1/(z-z_4)$ is equivalent to an inversion with respect to the point z_4, coupled with a certain reflexion (Ex. 21). If z_1, z_2, z_3 lie on a circle through z_4 the corresponding points $Z_1=1/(z_1-z_4)$, $Z_2=1/(z_2-z_4)$, $Z_3=1/(z_3-z_4)$ lie on a straight line. Hence (Ex. 12) we can find real quantities α', β', γ' such that $\alpha'+\beta'+\gamma'=0$ and $\alpha'/(z_1-z_4)+\beta'/(z_2-z_4)+\gamma'/(z_3-z_4)=0$, and it is easy to prove that this is equivalent to the given condition.

28. Prove the following analogue of De Moivre's Theorem for real quantities:—if ϕ_1, ϕ_2, ϕ_3, ... is a series of positive acute angles such that

$$\tan \phi_{m+1} = \tan \phi_m \sec \phi_1 + \sec \phi_m \tan \phi_1,$$

then

$$\tan \phi_{m+n} = \tan \phi_m \sec \phi_n + \sec \phi_m \tan \phi_n,$$

$$\sec \phi_{m+n} = \sec \phi_m \sec \phi_n + \tan \phi_m \tan \phi_n,$$

and

$$\tan \phi_m + \sec \phi_m = (\tan \phi_1 + \sec \phi_1)^m.$$

[Use the method of mathematical induction.]

29. **The transformation** $z = Z^m$. In this case $r = R^m$, and θ and $m\Theta$ differ by a multiple of 2π. If Z describes a circle round the origin, z describes a circle round the origin *m times*.

The whole (x, y) plane corresponds to any one of m sectors in the (X, Y) plane, each of angle $2\pi/m$. To each point in the (x, y) plane correspond m points in the (X, Y) plane.

30. **Complex functions of a real variable.** If $f(t)$, $\phi(t)$ are two real functions of a real variable t defined for a certain range of values of t, we call

$$z = f(t) + i\phi(t) \quad \text{..............................(1)}$$

a complex function of t. We can represent it graphically by drawing the curve

$$x = f(t), \quad y = \phi(t);$$

the equation of the curve may be obtained by eliminating t between these equations. If z is a polynomial in t, or rational function of t, with complex coefficients, we can express it in the form (1) and so determine the curve represented by the function.

(i) Let $$z = a + (b - a)t,$$
where a and b are complex numbers. If $a = \alpha + i\alpha'$, $b = \beta + i\beta'$, then

$$x = \alpha + (\beta - \alpha)t, \quad y = \alpha' + (\beta' - \alpha')t.$$

The curve is the straight line joining the points $z = a$ and $z = b$. The segment between the points corresponds to the range of values of t from 0 to 1. Find the values of t which correspond to the two produced segments of the line.

(ii) If $$z = c + \rho\{(1 + it)/(1 - it)\},$$
the curve is the circle of centre c and radius ρ. As t varies through all real values z describes the circle once.

(iii) In general the equation $z = (a + bt)/(c + dt)$ represents a circle. This can be proved by calculating x and y and eliminating: but this process is rather cumbrous. A simpler method is obtained by using the result of Ex. 22. Let $z = (a + bZ)/(c + dZ)$, $Z = t$. As t varies Z describes a straight line, viz. the axis of X. Hence z describes a circle.

(iv) The equation $$z = a + 2bt + ct^2$$
represents a parabola generally, a straight line if b/c is real.

(v) The equation $z=(a+2bt+ct^2)/(\alpha+2\beta t+\gamma t^2)$, where α, β, γ are real, represents a conic section.

[Eliminate t from

$$x=(A+2Bt+Ct^2)/(\alpha+2\beta t+\gamma t^2), \quad z=(A'+2B't+C't^2)/(\alpha+2\beta t+\gamma t^2)$$

(where $a=A+iA'$, $b=B+iB'$, $c=C+iC'$).]

38. Formulae for sin $n\theta$ and cos $n\theta$. De Moivre's Theorem enables us to express $\sin n\theta$ and $\cos n\theta$, where n is a positive integer, in terms of $\sin\theta$ and $\cos\theta$. For from the formula

$$\cos n\theta + i\sin n\theta = (\cos\theta + i\sin\theta)^n$$

(where n is a positive integer) we deduce

$$\cos n\theta = R\,[(\cos\theta + i\sin\theta)^n]$$
$$= (\cos\theta)^n \left\{1 - \binom{n}{2}\tan^2\theta + \binom{n}{4}\tan^4\theta - \dots\right\},$$
$$\sin n\theta = I\,[(\cos\theta + i\sin\theta)^n]$$
$$= (\cos\theta)^n \left\{\binom{n}{1}\tan\theta - \binom{n}{3}\tan^3\theta + \dots\right\},$$

where $\binom{n}{r}$ is the general binomial coefficient

$$\frac{n(n-1)(n-2)\dots(n-r+1)}{1.2.3\dots r}$$

(sometimes written ${}^nC_r$).

Then

$$\frac{\cos n\theta}{(\cos\theta)^n} = 1 - \binom{n}{2}t^2 + \binom{n}{4}t^4 - \dots \quad \dots\dots\dots\dots(1),$$
$$\frac{\sin n\theta}{(\cos\theta)^n} = \binom{n}{1}t - \binom{n}{3}t^3 + \dots \quad \dots\dots\dots\dots\dots\dots(2),$$

where $t = \tan\theta$.

By division

$$\tan n\theta = \left\{\binom{n}{1}t - \binom{n}{3}t^3 + \dots\right\}\Big/\left\{1 - \binom{n}{2}t^2 + \dots\right\} \quad \dots(3).$$

From (1) and (2) we can deduce further formulae expressing

$$\cos n\theta, \quad \frac{\sin n\theta}{\sin\theta}$$

in terms of $\cos\theta$ only. For

$$\cos^n\theta\tan^{2r}\theta = \sin^{2r}\theta\cos^{n-2r}\theta = \cos^{n-2r}\theta\,(1-\cos^2\theta)^r;$$

and on substituting in (1), after multiplying up by $\cos^n\theta$, we see that

$$\cos n\theta = a_n \cos^n\theta + a_{n-2}\cos^{n-2}\theta + \dots$$

where $a_n, a_{n-2}, \dots$ are constants, and the last term is a constant or a multiple of $\cos\theta$, according as n is even or odd.

Similarly, from (2), we deduce

$$\frac{\sin n\theta}{\sin\theta} = b_{n-1}\cos^{n-1}\theta + b_{n-3}\cos^{n-3}\theta + \dots .$$

To determine the actual values of the coefficients generally and directly, and by means of really elementary methods, is a matter of some little difficulty. The formulae are

$$2\cos n\theta = (2\cos\theta)^n - \frac{n}{1!}(2\cos\theta)^{n-2} + \frac{n(n-3)}{2!}(2\cos\theta)^{n-4}$$
$$- \dots + (-)^r \frac{n(n-r-1)\dots(n-2r+1)}{r!}(2\cos\theta)^{n-2r} + \dots \quad \dots(4),$$

$$\frac{\sin n\theta}{\sin\theta} = (2\cos\theta)^{n-1} - \frac{n-2}{1!}(2\cos\theta)^{n-3} + \frac{(n-3)(n-4)}{2!}(2\cos\theta)^{n-5}$$
$$- \dots + (-)^r \frac{(n-r-1)\dots(n-2r)}{r!}(2\cos\theta)^{n-2r-1} + \dots \quad \dots\dots(5).$$

That these formulae are correct is easily verified by induction. For

$$2\cos(n+1)\theta = \cos\theta(2\cos n\theta) - 2(1-\cos^2\theta)\frac{\sin n\theta}{\sin\theta},$$
$$\frac{\sin(n+1)\theta}{\sin\theta} = \cos\theta\left(\frac{\sin n\theta}{\sin\theta}\right) + \cos n\theta,$$

and if we assume that the formulae hold for $n = 1, 2, \dots k$ (and they are easily verified for $n = 1, 2, 3$), we can at once show that they hold for $n = k+1$. We leave this as an exercise for the reader.

39. When $\tan n\theta$ is given we can regard the equation (3), which we may write for brevity in the form $\tan n\theta = f(t)$, as an equation of the nth degree in t, one of whose roots is $t = \tan\theta$.

Similarly, one of the roots of

$$\tan\left\{n\left(\theta + \frac{k\pi}{n}\right)\right\} = f(t)$$

is $t = \tan\{\theta + (k\pi/n)\}$ (k being any integer). But since

$$\tan(n\theta + k\pi) = \tan n\theta,$$

the two equations are the same. Hence $\tan\{\theta+(k\pi/n)\}$ is a root of (3) for all values of k.

As we give k all integral values, this expression assumes n and only n distinct values, viz.

$$\tan\theta,\ \tan\left(\theta+\frac{\pi}{n}\right),\ \tan\left(\theta+\frac{2\pi}{n}\right),\ \ldots,\ \ldots\ \tan\left\{\theta+\frac{(n-1)\pi}{n}\right\}.$$

These are therefore the roots of (3), considered as an equation in t. It follows that any symmetric function of these quantities can be expressed in terms of the coefficients of (3), i.e. in terms of $\tan n\theta$. The equations (1), (2), (4), (5) can of course be considered from the same point of view. Some illustrations will be found in the examples which follow.

Examples XXIV. 1. The equation (3) may be written in the form

$$t^n-\binom{n}{2}t^{n-2}+\binom{n}{4}t^{n-4}-\ldots-\frac{\tan}{\cot}n\theta\left\{\binom{n}{1}t^{n-1}-\binom{n}{3}t^{n-3}+\ldots\right\}=0,$$

where $\tan n\theta$ or $\cot n\theta$ is to be chosen according as n is odd or even.

2. Show that

$$\sec^2\theta+\sec^2\left(\theta+\frac{\pi}{n}\right)+\ldots+\sec^2\left\{\theta+\frac{(n-1)\pi}{n}\right\}$$

is equal to $n^2\sec^2 n\theta$ or $n^2\operatorname{cosec}^2 n\theta$ according as n is odd or even.

(*Math. Trip.* 1900.)

[The expression given is $n+\Sigma t_r^2$, where t_r is a root of the equation in Ex. 1.]

3. Prove that $\sec^2\frac{\pi}{9}+\sec^4\frac{2\pi}{9}+\sec^4\frac{3\pi}{9}+\sec^4\frac{4\pi}{9}=1120.$

4. If n is odd

$$t^n-\binom{n}{2}t^{n-2}+\binom{n}{4}t^{n-4}-\ldots\equiv t\left(t^2-\tan^2\frac{\pi}{n}\right)\left(t^2-\tan^2\frac{2\pi}{n}\right)\ldots\left(t^2-\tan^2\frac{r\pi}{n}\right),$$

where $r=\frac{1}{2}(n-1)$. State and prove the corresponding formula when n is even.

5. The roots of the equation $2\cos n\theta=x^n-\frac{n}{1!}x^{n-2}+\frac{n(n-3)}{2!}x^{n-4}-\ldots$ are

$$2\cos\theta,\quad 2\cos\left(\theta+\frac{2\pi}{n}\right),\ \ldots\quad 2\cos\left\{\theta+\frac{2(n-1)\pi}{n}\right\}.$$

6. The roots of $64x^3-112x^2+56x-7=0$ are $\sin^2\frac{\pi}{7}$, $\sin^2\frac{2\pi}{7}$, $\sin^2\frac{4\pi}{7}$.

Deduce that $$\sin\frac{2\pi}{7}+\sin\frac{4\pi}{7}-\sin\frac{\pi}{7}=\tfrac{1}{2}\sqrt{7}.$$

7. Show that $4\cos^2(\pi/7)$ is a root of $x^3-5x^2+6x-1=0$, and find the other roots. (*Math. Trip.* 1898.)

8. **Factor-formulae for** $\cos n\theta$ **and** $(\sin n\theta)/(\sin\theta)$*. Show that $2\cos n\theta$ is equal to

$$2^n\left(\cos^2\theta-\cos^2\frac{\pi}{2n}\right)\left(\cos^2\theta-\cos^2\frac{3\pi}{2n}\right)\ldots\ldots\left\{\cos^2\theta-\cos^2\frac{(n-2)\pi}{2n}\right\}\cos\theta,$$

or to $$2^n\left(\cos^2\theta-\cos^2\frac{\pi}{2n}\right)\left(\cos^2\theta-\cos^2\frac{3\pi}{2n}\right)\ldots\ldots\left\{\cos^2\theta-\cos^2\frac{(n-1)\pi}{2n}\right\},$$

according as n is odd or even.

9. Show that $$\sin\frac{\pi}{2n}\sin\frac{3\pi}{2n}\ldots\sin\frac{r\pi}{2n}=2^{-\frac{1}{2}(n-1)},$$

where r is equal to $n-2$ or to $n-1$ according as n is odd or even.

[Put $\theta=0$ in Ex. 8: consider the sign carefully, when taking the square root of each side.]

10. Show that $(\sin n\theta)/(\sin\theta)$ is equal to

$$2^{n-1}\left(\cos^2\theta-\cos^2\frac{\pi}{n}\right)\left(\cos^2\theta-\cos^2\frac{2\pi}{n}\right)\ldots\ldots\left\{\cos^2\theta-\cos^2\frac{(n-2)\pi}{2n}\right\}\cos\theta,$$

or to $$2^{n-1}\left(\cos^2\theta-\cos^2\frac{\pi}{n}\right)\left(\cos^2\theta-\cos^2\frac{2\pi}{n}\right)\ldots\ldots\left\{\cos^2\theta-\cos^2\frac{(n-1)\pi}{2n}\right\},$$

according as n is even or odd.

11. Show from Ex. 10, and equation (2) of § 38, that if n is odd

$$2^{n-1}\left(\cos^2\theta-\cos^2\frac{\pi}{n}\right)\left(\cos^2\theta-\cos^2\frac{2\pi}{n}\right)\ldots\left\{\cos^2\theta-\cos^2\frac{(n-1)\pi}{2n}\right\}$$
$$=n\cos^{n-1}\theta-\frac{n(n-1)(n-2)}{3!}\cos^{n-3}\theta\sin^2\theta+\ldots,$$

and deduce, by putting $\theta=0$, that

$$\sin\frac{\pi}{n}\sin\frac{2\pi}{n}\ldots\sin\frac{(n-1)\pi}{2n}=2^{-\frac{1}{2}(n-1)}\sqrt{n}.$$

Obtain the corresponding formula for the case in which n is even.

40. Roots of complex numbers. We have not, up to the present, attributed any meaning to symbols such as $\sqrt[n]{a}$, $a^{m/n}$, when a is a complex number, and m and n integers. It is, however, natural to adopt the definitions which are given in elementary algebra for real values of a. Thus we define $\sqrt[n]{a}$ or $a^{1/n}$, where n is a positive integer, as any number a which satisfies the equation $z^n=a$; and $a^{m/n}$, where m is an integer, as $(a^{1/n})^m$. These definitions do not prejudge the question as to whether there are or are not more than one (or any) roots of the equation.

41. Solution of the equation $z^n=a$. Let

$$a=\rho(\cos\phi+i\sin\phi),$$

where ρ is positive and ϕ is an angle such that $-\pi<\phi\leqq\pi$.

* The results of Exs. 8—11 are of considerable importance in Higher Trigonometry.

Then if $z = r(\cos\theta + i\sin\theta)$, the equation takes the form

$$r^n(\cos n\theta + i\sin n\theta) = \rho(\cos\phi + i\sin\phi),$$

so that $$r^n = \rho, \quad \cos n\theta = \cos\phi, \quad \sin n\theta = \sin\phi \quad\text{.........(1).}$$

The only possible value of r is $\sqrt[n]{\rho}$, the ordinary arithmetical nth-root of ρ; and in order that the last two equations should be satisfied it is necessary and sufficient that $n\theta = \phi + 2k\pi$, where k is an integer, or

$$\theta = (\phi + 2k\pi)/n.$$

If $k = pn + q$, where p and q are integers, and $0 \leqq q < n$, the value of θ is $2p\pi + (\phi + 2q\pi)/n$, and in this the value of p is a matter of indifference. Hence *the equation*

$$z^n = a = \rho(\cos\phi + i\sin\phi)$$

has n roots and n only, given by $z = r(\cos\theta + i\sin\theta)$, *where*

$$r = \sqrt[n]{\rho}, \quad \theta = (\phi + 2q\pi)/n \quad (q = 0, 1, 2, \dots n-1).$$

That these n roots are in reality all distinct is easily seen by plotting them on the Argand Diagram. The figure (Fig. 34) shows the four fourth roots of

$$(1{\cdot}6)(\cos 55^\circ + i\sin 55^\circ).$$

The particular root

$$\sqrt[n]{\rho}\,\{\cos(\phi/n) + \sin(\phi/n)\}$$

is called the *principal value* of $\sqrt[n]{a}$.

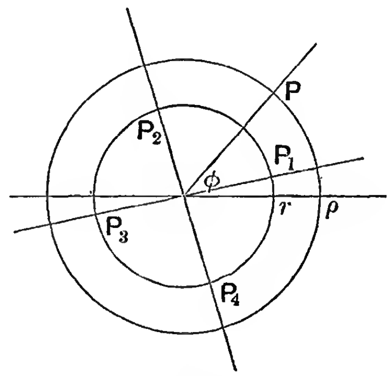


FIG. 34.

The case in which $a = 1$, $\rho = 1$, $\phi = 0$ is of particular interest. The n roots of the equation $z^n = 1$ are

$$\cos(2q\pi/n) + i\sin(2q\pi/n), \quad (q = 0, 1, \dots n-1).$$

These quantities are called the nth roots of unity; the principal value is unity itself. If we write ω_n for $\cos(2\pi/n) + i\sin(2\pi/n)$ we see that the n roots of unity are

$$1, \ \omega_n, \ \omega_n^2, \dots \omega_n^{n-1}.$$

Examples XXV. 1. The two square roots of 1 are $1, -1$; the three cube roots are 1, $\frac{1}{2}(-1+i\sqrt{3})$, $\frac{1}{2}(-1-i\sqrt{3})$; the four fourth roots are 1, i, -1, $-i$; and the five fifth roots are

$$1, \ \tfrac{1}{4}\{\sqrt{5}-1+i\sqrt{10+2\sqrt{5}}\}, \ \tfrac{1}{4}\{-\sqrt{5}-1+i\sqrt{10-2\sqrt{5}}\},$$
$$\tfrac{1}{4}\{-\sqrt{5}-1-i\sqrt{10-2\sqrt{5}}\}, \ \tfrac{1}{4}\{\sqrt{5}-1-i\sqrt{10+2\sqrt{5}}\}.$$

2. Prove that $$1+\omega_n+\omega_n^2+\dots+\omega_n^{n-1}=0.$$

3. Prove that $(x+y\omega_3+z\omega_3^2)(x+y\omega_3^2+z\omega_3)=x^2+y^2+z^2-yz-zx-xy$.

4. The nth roots of a are $\sqrt[n]{a}$, $\omega_n\sqrt[n]{a}$, $\omega_n^2\sqrt[n]{a}$, $\dots\ \omega_n^{n-1}\sqrt[n]{a}$, where $\sqrt[n]{a}$ denotes the *principal* nth root.

5. It follows from Exs. XXIII. 14 that the roots of
$$z^2=\alpha+i\beta$$
are
$$\pm\sqrt{[\tfrac{1}{2}\{\sqrt{(\alpha^2+\beta^2)}+\alpha\}]}\pm i\sqrt{[\tfrac{1}{2}\{\sqrt{(\alpha^2+\beta^2)}-\alpha\}]},$$
like or unlike signs being chosen according as β is positive or negative. Show that this result agrees with the result of § 41.

6. Show that $(x^{2m}-a^{2m})/(x^2-a^2)$ is equal to
$$\left(x^2-2ax\cos\frac{\pi}{m}+a^2\right)\left(x^2-2ax\cos\frac{2\pi}{m}+a^2\right)\dots\left(x^2-2ax\cos\frac{(m-1)\pi}{m}+a^2\right).$$

[The factors of $x^{2m}-a^{2m}$ are
$$(x-a),\quad (x-a\omega_{2m}),\quad (x-a\omega^2_{2m}),\ \dots(x-a\omega_{2m}^{2m-1}).$$
The factor $x-a\omega_{2m}{}^m$ is $x+a$. The factors $(x-a\omega_{2m}{}^s)$, $(x-a\omega_{2m}^{2m-s})$ taken together give a factor $x^2-2ax\cos\dfrac{s\pi}{m}+a^2$.]

7. Resolve $x^{2m+1}-a^{2m+1}$, $x^{2m}+a^{2m}$, and $x^{2m+1}+a^{2m+1}$ into factors in a similar way.

8. Show that $x^{2n}-2x^na^n\cos\theta+a^{2n}$ is equal to
$$\left(x^2-2xa\cos\frac{\theta}{n}+a^2\right)\left(x^2-2xa\cos\frac{\theta+2\pi}{n}+a^2\right)\dots$$
$$\dots\left(x^2-2xa\cos\frac{\theta+2(n-1)\pi}{n}+a^2\right).$$

[Use the formula
$$x^{2n}-2x^na^n\cos\theta+a^{2n}=\{x^n-a^n(\cos\theta+i\sin\theta)\}\{x^n-a^n(\cos\theta-i\sin\theta)\},$$
and split up each of the last two expressions into n factors.]

9. The problem of finding the accurate value of ω_n in a numerical form involving only square roots, as in the formula $\omega_3=\frac{1}{2}(-1+i\sqrt{3})$, is the algebraical equivalent of the geometrical problem of inscribing in a circle a regular polygon of n sides by Euclidean methods, i.e. by ruler and compasses. We saw in fact in Chapters I. and II. that irrationals involving square roots could always be so constructed, and are the only irrationals which can be so constructed.

Euclid gives constructions for $n=3$, 4, 5, 6, 8, 10, 12, and 15. It is evident that the construction is possible for any value of n which can be found from these by multiplication by any power of 2. There are other special values of n for which such constructions are possible, the most interesting being $n=17$.

Approximate constructions for regular polygons of any number of sides will be found in books of practical geometry.

42. The general form of De Moivre's Theorem. It follows from the results of the last section that, if q is a positive integer, one of the values of $(\cos\theta + i\sin\theta)^{1/q}$ is

$$\cos(\theta/q) + i\sin(\theta/q).$$

Raising each of these expressions to the power p (where p is any integer positive or negative), we obtain the theorem that one of the values of $(\cos\theta + i\sin\theta)^{p/q}$ is $\cos(p\theta/q) + i\sin(p\theta/q)$, or *if α is any rational quantity, one of the values of* $(\cos\theta + i\sin\theta)^{\alpha}$ *is*

$$\cos\alpha\theta + i\sin\alpha\theta.$$

This is a generalised form of De Moivre's theorem stated in § 36.

MISCELLANEOUS EXAMPLES ON CHAPTER III.

1. The condition that a triangle (xyz) should be equilateral is that

$$x^2+y^2+z^2-yz-zx-xy=0.$$

[Let XYZ be the triangle. The displacement $\overline{ZX}$ is $\overline{YZ}$ turned through an angle $\frac{2}{3}\pi$ in the positive or negative direction: or, as $\operatorname{Cis}\frac{2}{3}\pi=\omega_3$, $\operatorname{Cis}(-\frac{2}{3}\pi)=1/\omega_3=\omega_3^2$, we have $x-z=(z-y)\omega_3$ or $x-z=(z-y)\omega_3^2$. Hence $x+y\omega_3+z\omega_3^2=0$ or $x+y\omega_3^2+z\omega_3=0$. The result follows from Ex. XXV. 3.]

2. If XYZ, $X'Y'Z'$ are two triangles, and

$$\overline{YZ}\,.\,\overline{Y'Z'}=\overline{ZX}\,.\,\overline{Z'X'}=\overline{XY}\,.\,\overline{X'Y'},$$

then both triangles are equilateral.

[From the equations

$$(y-z)(y'-z')=(z-x)(z'-x')=(x-y)(x'-y')=\kappa^2$$

say, we deduce $\Sigma\, 1/(y'-z')=0$, or $\Sigma x'^2-\Sigma y'z'=0$. Now apply the last example.]

3. On the sides of a triangle ABC similar triangles BCX, CAY, ABZ are described. Show that the centres of gravity of ABC, XYZ are coincident.

[We have $(x-c)/(b-c)=(y-a)/(c-a)=(z-b)/(a-b)=\lambda$, say. Express $\frac{1}{3}(x+y+z)$ in terms of a, b, c.]

4. If X, Y, Z are points on the sides of the triangle ABC, such that

$$BX/XC=CY/YA=AZ/ZB=r,$$

and if ABC, XYZ are similar, then either $r=1$ or both triangles are equilateral.

5. Deduce Ptolemy's Theorem concerning cyclic quadrilaterals from the fact that the cross ratios of four concyclic points are real.

[Start from the identity

$$(x_2-x_3)(x_1-x_4)+(x_3-x_1)(x_2-x_4)+(x_1-x_2)(x_3-x_4)=0.]$$

6. If $z^2+z'^2=1$, the points z, z' are ends of conjugate diameters of an ellipse whose foci are the points $1, -1$. [If CP, CD are conjugate semi-diameters of an ellipse and S, H its foci, then CD is parallel to the exterior bisector of the angle SPH, and $SP \,.\, HP=CD^2$.]

7. Prove that $|a+b|^2+|a-b|^2=2\{|a|^2+|b|^2\}$. [This is the analytical equivalent of the geometrical theorem that, if M is the middle point of PQ, $OP^2+OQ^2=2OM^2+2MP^2$.]

8. Deduce from Ex. 7 that

$$|a+\sqrt{(a^2-b^2)}|+|a-\sqrt{(a^2-b^2)}|=|a+b|+|a-b|.$$

[If $a+\sqrt{(a^2-b^2)}=z_1$, $a-\sqrt{(a^2-b^2)}=z_2$, we have

$$|z_1|^2+|z_2|^2=\tfrac{1}{2}|z_1+z_2|^2+\tfrac{1}{2}|z_1-z_2|^2=2|a|^2+2|a^2-b^2|,$$

and so $(|z_1|+|z_2|)^2=2\{|a|^2+|a^2-b^2|+|b|^2\}=|a+b|^2+|a-b|^2+2|a^2-b^2|$.

Another way of stating the result is: if z_1 and z_2 are the roots of $\alpha z^2+2\beta z+\gamma=0$, then

$$|z_1|+|z_2|=(1/|\alpha|)\{(|-\beta+\sqrt{\alpha\gamma}|)+(|-\beta-\sqrt{\alpha\gamma}|)\}.]$$

9. If $x^4+4a_1x^3+6a_2x^2+4a_3x+a_4=0$ is an equation with real coefficients and has two real and two complex roots, concyclic in the Argand diagram, then

$$a_3^2+a_1^2a_4+a_2^3-a_2a_4-2a_1a_2a_3=0.$$

10. The four roots of $a_0x^4+4a_1x^3+6a_2x^2+4a_3x+a_4=0$ will be harmonically related if

$$a_0a_3^2+a_1^2a_4+a_2^3-a_0a_2a_4-2a_1a_2a_3=0.$$

[Express $Z_{23,14}Z_{31,24}Z_{12,34}$, where $Z_{23,14}=(z_1-z_2)(z_3-z_4)+(z_1-z_3)(z_2-z_4)$ and z_1, z_2, z_3, z_4 are the roots of the equation, in terms of the coefficients.]

11. **Imaginary points and straight lines.** Let $ax+by+c=0$ be an equation with complex coefficients (which of course may be real in special cases).

If we give x any particular real or complex value, we can find the corresponding value of y. The aggregate of pairs of real or complex values of x and y which satisfy the equation is called an *imaginary straight line*; the pairs of values are called *imaginary points*, and are said *to lie on the line*. The values of x and y are called the *coordinates* of the point (x, y). When x and y are real, the point is called a *real point*: when a, b, c are all real (or can be made all real by division by a factor), the line is called a *real line*. The points $x=\alpha+i\beta$, $y=\gamma+i\delta$ and $x=\alpha-i\beta$, $y=\gamma-i\delta$ are said to be *conjugate*; and so are the lines

$$(A+iA')x+(B+iB')y+C+iC'=0, \quad (A-iA')x+(B-iB')y+C-iC'=0.$$

Verify the following assertions:—every real line contains infinitely many pairs of conjugate imaginary points; every imaginary line contains one and only one real point; an imaginary line cannot contain a pair of conjugate imaginary points:—and find the conditions (a) that the line joining two given imaginary points should be real, and (b) that the point of intersection of two imaginary lines should be real.

12. Sum the series

$$\cos a + \cos(a+b) + \cos(a+2b) + \dots, \quad \sin a + \sin(a+b) + \sin(a+2b) + \dots$$

to n terms.

[Sum the geometrical series

$$(\cos a + i\sin a)\{1 + (\cos b + i\sin b) + (\cos b + i\sin b)^2 + \dots\}$$

to n terms, and equate real and imaginary parts.]

13. Also sum

$$1 + \binom{n}{1}\cos a + \binom{n}{2}\cos 2a + \dots, \quad 1 + \binom{n}{1}\sin a + \binom{n}{2}\sin 2a + \dots$$

to $n+1$ terms.

14. Sum the series $\cos a + x\cos(a+\beta) + \dots + x^{n-1}\cos\{a+(n-1)\beta\}$.

(*Math. Trip.* 1905.)

15. Find the modulus and the amplitude of

$$1+\cos\theta+i\sin\theta, \quad 1+\cos\theta-i\sin\theta, \quad 1-\cos\theta+i\sin\theta, \quad 1-\cos\theta-i\sin\theta.$$

16. Find the square roots of the numbers in the preceding example.

17. Prove that

$$\left(\frac{1+\sin\theta+i\cos\theta}{1+\sin\theta-i\cos\theta}\right)^n = \cos\left(\tfrac{1}{2}n\pi - n\theta\right) + i\sin\left(\tfrac{1}{2}n\pi - n\theta\right).$$

18. Prove the identities

$$(x+y+z)(x+\omega_3 y+\omega_3^2 z)(x+\omega_3^2 y+\omega_3 z) = x^3+y^3+z^3-3xyz,$$

$$(x+y+z)(x+\omega_5 y+\omega_5^4 z)(x+\omega_5^2 y+\omega_5^3 z)(x+\omega_5^3 y+\omega_5^2 z)(x+\omega_5^4 y+\omega_5 z)$$
$$= x^5+y^5+z^5-5x^3yz+5xy^2z^2.$$

19. Solve the equations

$$x^3-3ax+(a^3+1)=0 \text{ and } x^5-5ax^3+5a^2x+(a^5+1)=0.$$

20. If $f(x)=a_0+a_1x+\dots+a_kx^k$, then

$$\{f(x)+f(\omega x)+\dots+f(\omega^{n-1}x)\}/n = a_0+a_nx^n+a_{2n}x^{2n}+\dots+a_{\lambda n}x^{\lambda n},$$

ω being any root of $x^n=1$ (except $x=1$), and λn the greatest multiple of n contained in λ. Find a similar formula for $a_\mu + a_{\mu+n}x^n + a_{\mu+2n}x^{2n} + \dots$.

21. If $$(1+x)^n = p_0+p_1x+p_2x^2+\dots$$

(n being a positive integer), prove that

$$p_0-p_2+p_4-\dots = 2^{\frac{1}{2}n}\cos\tfrac{1}{4}n\pi, \quad p_1-p_3+p_5-\dots = 2^{\frac{1}{2}n}\sin\tfrac{1}{4}n\pi.$$

22. Sum the series

$$\frac{x}{2!\,n-2!} + \frac{x^2}{5!\,n-5!} + \frac{x^3}{8!\,n-8!} + \dots + \frac{x^{n/3}}{n-1!},$$

n being a multiple of 3. (*Math. Trip.* 1899.)

23. Show how to deduce the formulae given in § 38 from the addition theorems for $\cos x$ and $\sin x$, using no complex quantities. [See Hobson's *Trigonometry*, Ch. VII.]

24. Prove that, when m is odd, $3\Sigma \operatorname{cosec}^2 (r\pi/m) = m^2 - 1$, the summation extending to $r = 1, 2, \ldots, m-1$.

The corresponding sum, extended only to values of r less than m and prime to it, is denoted by S. Show that, if m is the product of unequal odd primes $a, b, c, \ldots k$, then $3S = (a^2-1)(b^2-1)(c^2-1)\ldots(k^2-1)$.

(*Math. Trip.* 1902.)

25. Prove that, if m is odd,

$$\sum_{r=0}^{m-1} \tan^4 \{(2r+1)\pi/m\} = \tfrac{1}{3} m(m-1)(m^2+m-3).$$

(*Math. Trip.* 1903.)

26. If $a_r = a + b\cos\{\theta + (2r\pi/n)\}$, for $r = 1, 2, \ldots, n$, show that

$$3n(n-2)\Sigma a_1 \Sigma a_1 a_2 = 3n^2 \Sigma a_1 a_2 a_3 + (n-1)(n-2)(\Sigma a_1)^3,$$

if $n > 3$, and find the corresponding equation when $n = 3$. (*Math. Trip.* 1906.)

27. The roots of

$$1 - nx - \frac{n(n-1)}{2!}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \ldots + (-1)^{\frac{1}{2}n(n+1)}x^n = 0$$

are

$$\tan\frac{\pi}{4n},\quad \tan\frac{5\pi}{4n},\ \ldots,\ \tan\frac{(4n-3)\pi}{4n}.$$

[The equation is $(1-xi)^n + i(1+xi)^n = 0$.]

28. If $a = \pi/4n$, show (cf. Ex. 27) that

$$\cot a,\quad -\cot 3a,\quad \cot 5a,\ \ldots,\ (-)^{n-1}\cot(2n-1)a$$

are the roots of

$$x^n - nx^{n-1} - \frac{n(n-1)}{2!}x^{n-2} + \frac{n(n-1)(n-2)}{3!}x^{n-3} - \ldots = 0.$$

Deduce that

$$\cot a \operatorname{cosec}^2 a - \cot 3a \operatorname{cosec}^2 3a + \ldots \text{ to } n \text{ terms}$$

is equal to $2n^3$. (*Math. Trip.* 1901.)

29. There are in general two points unaltered by the transformation $z = (aZ+b)/(cZ+d)$. If these points are α, β, the transformation can be put in the form $(z-\alpha)/(z-\beta) = k(Z-\alpha)/(Z-\beta)$.

In particular, reduce the transformation $z = (1+Z)/(1-Z)$ to this form. Divide the Z-plane into 8 regions by means of the axes and the unit circle. Find the region in the z-plane which corresponds to each of these regions.

30. If $z = Z^2 - 1$, then as z describes the circle $|z| = \kappa$, the two corresponding positions of Z each describe the Cassinian oval $\rho_1\rho_2 = \kappa$, where ρ_1, ρ_2 are the distances from the points ± 1. Trace the ovals for different values of κ.

31. If t is a complex number such that $|t| = 1$, then as t varies, the point $x = (at+b)/(t-c)$ describes a circle, unless $|c| = 1$, when it describes a straight line.

32. If t varies as in the last example, the point $x=\frac{1}{2}\{at+(b/t)\}$ in general describes an ellipse whose foci are given by $x^2=ab$, and whose axes are $|a|+|b|$ and $|a|-|b|$. But if $|a|=|b|$, x describes the finite straight line joining the points $\pm\sqrt{(ab)}$.

33. If $z=2Z+Z^2$, the circle $|Z|=1$ corresponds to a cardioid in the plane of z.

34. Discuss the transformation $z=\frac{1}{2}\{Z+(1/Z)\}$, showing in particular that to the circles $X^2+Y^2=a^2$ correspond the confocal ellipses

$$\frac{x^2}{\left\{\frac{1}{2}\left(a+\frac{1}{a}\right)\right\}^2}+\frac{y^2}{\left\{\frac{1}{2}\left(a-\frac{1}{a}\right)\right\}^2}=1.$$

35. If $(z+1)^2=4/Z$, the unit circle in the z-plane corresponds to the parabola $R\cos^2\frac{1}{2}\Theta=1$ in the Z-plane, and the inside of the circle to the outside of the parabola.

36. Show that, by means of the transformation $z=\{(Z-ic)/(Z+ic)\}^2$, the upper half of the z-plane may be made to correspond to the interior of a certain semicircle in the Z-plane.

37. Consider the relation $az^2+2hzZ+bZ^2+2gz+2fZ+c=0$.

Show that there are two values of Z for which the corresponding values of z are equal, and *vice versa*. We call these the *branch points* in the Z and z-planes respectively. Show that, if z describes an ellipse whose foci are the branch points, so does Z.

[We can, without loss of generality, take the given relation in the form

$$z^2+2zZ\cos\omega+Z^2=1$$

—the reader should satisfy himself that this is the case. The branch points in either plane are $\pm\operatorname{cosec}\omega$. An ellipse of the form specified is given by

$$|z+\operatorname{cosec}\omega|+|z-\operatorname{cosec}\omega|=\text{const.}$$

This is equivalent (Ex. 8) to

$$|z+\sqrt{(z^2-\operatorname{cosec}^2\omega)}|+|z-\sqrt{(z^2-\operatorname{cosec}^2\omega)}|=\text{const.}$$

Express this in terms of Z.]

38. If $z=aZ^m+bZ^n$, where m, n are positive integers and a, b real, show that as Z describes the unit circle, z describes a hypo- or epi-cycloid.

39. Prove that

$$\frac{\sin(2n+1)\theta}{(2n+1)\sin\theta}=\prod_{r=1}^{n}\left(1-\frac{\sin^2\theta}{\sin^2\{r\pi/(2n+1)\}}\right). \quad (\textit{Math. Trip.}\ 1907.)$$

40. By putting $\theta=\frac{1}{2}\pi$ in the last example, prove that

$$\cot\frac{\pi}{2n+1}\cot\frac{2\pi}{2n+1}\cdots\cot\frac{n\pi}{2n+1}=\frac{1}{\sqrt{(2n+1)}}.$$

41. Prove that

$$\sin n\theta = 2^{n-1} \sin\theta \sin\left(\theta + \frac{\pi}{n}\right) \dots \sin\left\{\theta + \frac{(n-1)\pi}{n}\right\}.$$

[Put $x=a=1$ in Ex. XXV. 8, and change θ into 2θ.]

42. Prove that $\quad \cos\frac{\pi}{15}\cos\frac{2\pi}{15}\dots\cos\frac{7\pi}{15} = \frac{1}{128}.$

43. Prove that $\quad \tan\frac{\pi}{2n}\tan\frac{2\pi}{2n}\dots\tan\frac{(n-1)\pi}{2n} = 1.$

44. Prove that

$$\frac{(1+x)^n - (1-x)^n}{2x} = A\left(x^2 + \tan^2\frac{\pi}{n}\right)\left(x^2 + \tan^2\frac{2\pi}{n}\right)\dots\left(x^2 + \tan^2\frac{r\pi}{n}\right),$$

where $A=1$, $r=\frac{1}{2}(n-1)$ if n is odd, and $A=n$, $r=\frac{1}{2}n-1$ if n is even.

45. If $1\Big/\prod_{r=1}^{n}\left(x+\tan^2\frac{r\pi}{2n+1}\right)$ is expressed in the form

$$\sum_{r=1}^{n} A_r\Big/\left(x+\tan^2\frac{r\pi}{2n+1}\right),$$

n being a positive integer, show that

$$A_r = \frac{(-1)^{r-1}2}{2n+1}\sin^2\frac{r\pi}{2n+1}\cos^{2n-3}\frac{r\pi}{2n+1}.$$

(*Math. Trip.* 1905.)

[Apply the ordinary rule for partial fractions: it will be found that

$$A_r = (-1)^{r-1}\,2\sin^2\frac{r\pi}{2n+1}\cos^{2n-3}\frac{r\pi}{2n+1}\prod_{k=1}^{n}\cot^2\frac{k\pi}{2n+1},$$

and Ex. 40 can be used to obtain the given result.]

46. Show that

$$\sum_{r=0}^{n-1}\sin\frac{(2r+1)\pi}{2n}\operatorname{cosec}\left\{\frac{(2r+1)\pi}{2n} - a\right\} = n\cos(n-1)a\sec na.$$

(*Math. Trip.* 1907.)

[The right-hand side is

$$n\frac{x^{n-1}+x^{-(n-1)}}{x^n+x^{-n}} = n\frac{x^{2n-1}+x}{x^{2n}+1},$$

where $x=\cos a + i\sin a = \operatorname{Cis} a$. The roots of $x^{2n}+1=0$ are

$$\operatorname{Cis}\frac{(2r+1)\pi}{2n}\quad (r=0,\ 1,\ \dots,\ 2n-1).$$

Split up the right-hand side into partial fractions of the form

$$A_r\Big/\left\{x - \operatorname{Cis}\frac{(2r+1)\pi}{2n}\right\}.$$

It will be found that $A_r = -i\sin\frac{(2r+1)\pi}{2n}\operatorname{Cis}\frac{(2r+1)\pi}{2n}$. To get the result in the form given we must associate the terms in pairs $(r,\ n+r)$ where $r=0,\ 1,\ \dots,\ n-1$.]

47. Show that, if m and n are positive integers, and $m \leqq n$, then $x^{m-1}/(1+x^n)=(1/n)\Sigma\, \omega^{m-n}/(x-\omega)$, where ω is a root of $x^n+1=0$: and hence show that, if n is even,

$$\frac{1}{1+x^n}=\frac{2}{n}\sum_0^{\frac{1}{2}n-1}\left\{1-x\cos\frac{(2r+1)\pi}{n}\right\}\bigg/\left\{1-2x\cos\frac{(2r+1)\pi}{n}+x^2\right\},$$

and find the corresponding formula when n is odd.

48. Express $x^{m-1}/(1-x^n)$, where m and n are positive integers, and $m \leqq n$, in partial fractions, and obtain the formulae for $1/(1-x^n)$ corresponding to those of Ex. 47.

49. Show that

$$\frac{x^n-a^n\cos n\theta}{x^{2n}-2x^n a^n\cos n\theta+a^{2n}}=\frac{1}{nx^{n-1}}\sum_{r=0}^{n-1}\frac{x-a\cos\left(\theta+\dfrac{2r\pi}{n}\right)}{x^2-2xa\cos\left(\theta+\dfrac{2r\pi}{n}\right)+a^2}.$$

50. If $\rho_1, \rho_2, \ldots \rho_n$ are the distances of a point P, in the plane of a regular polygon, from the vertices, prove that

$$\sum_1^n\frac{1}{\rho_\nu^2}=\frac{n}{r^2-a^2}\,\frac{r^{2n}-a^{2n}}{r^{2n}-2r^n a^n\cos n\theta+a^{2n}},$$

where O is the centre and a the radius of the circumcircle of the polygon, r the length OP, and θ the angle between OP and the radius from O to any vertex of the polygon.

51. If $A_1A_2\ldots A_n$, $B_1B_2\ldots B_m$ are concentric regular polygons, m and n being prime to one another, prove that

$$\sum_{r=1}^{n}\sum_{s=1}^{m}\frac{1}{(A_rB_s)^2}=\frac{mn}{b^2-a^2}\,\frac{b^{2mn}-a^{2mn}}{b^{2mn}-2b^{mn}a^{mn}\cos^{mn}\theta+a^{2mn}},$$

where a and b are the radii of the circumcircles of the polygons, and θ the angle between any two radii drawn one to a vertex of each polygon.

(*Math. Trip.* 1903.)

52. If p and q are integers, and q prime to p, and k is an odd positive integer less than $2p$, and $\theta=q\pi/p$, show that

$$\sum_{n=0}^{p-1}\frac{\cos k\,(a+n\theta)}{\sin\,(a+n\theta)}=p\cot pa,\qquad \sum_{n=0}^{p-1}\frac{\sin k\,(a+n\theta)}{\sin\,(a+n\theta)}=p.$$

[We have

$$\frac{px^{\lambda-1}}{x^p-1}-\frac{1}{x-1}=\sum_1^{p-1}\frac{t^{n\lambda}}{x-t^n},$$

where

$$t=\cos 2\theta+i\sin 2\theta,\quad 1\leqq\lambda\leqq p.$$

In this equation write $\frac{1}{2}(k+1)$ for λ and $\cos 2a-i\sin 2a$ for x.]

CHAPTER IV.

LIMITS OF FUNCTIONS OF A POSITIVE INTEGRAL VARIABLE.

43. Functions of a positive integral variable. In Chapter II. we discussed the notion of a function of a real variable x, and illustrated the discussion by a large number of examples of such functions. And the reader will remember that there was one important particular with regard to which the functions which we took as illustrations differed very widely. Some were defined for *all* values of x, some for *rational* values only, some for *integral* values only, and so on.

Consider, for example, the following functions: (i) $y=x$, (ii) $y=\sqrt{x}$, (iii) $y=$ the denominator of x, (iv) $y=$ the square root of the product of the numerator and the denominator of x, (v) $y=$ the largest prime factor of x, (vi) $y=$ the product of $\sqrt{x}$ and the largest prime factor of x, (vii) $y=$ the xth prime number, (viii) $y=$ the height measured in inches of convict x in Dartmoor prison.

Then the aggregates of values of x for which these functions are defined or, as we may say, the *fields of definition* of the functions, consist of (i) *all* values of x, (ii) *all positive* values of x, (iii) *all rational* values of x, (iv) *all positive rational* values of x, (v) *all integral* values of x, (vi), (vii) *all positive integral* values of x, (viii) a certain number of positive integral values of x, viz., $1, 2, \dots, N$, where N is the total number of convicts at Dartmoor at the present moment of time*.

Now let us consider a function, such as (vii) above, which is defined for all positive integral values of x and no others. This function may be regarded from two slightly different points of

* In the last case N depends on the time, and convict x, where x has a definite value, is a different individual at different moments of time. Thus if we take different moments of time into consideration we have a simple example of a function $y=F(x, t)$ of *two* variables, defined for a certain range of values of t, viz. from the time of the establishment of Dartmoor prison to the time of its abandonment, and for a certain number of positive integral values of x, this number varying with t.

view. We may consider it, as has so far been our custom, as a function of the real variable x defined for some only of the values of x, viz. positive integral values, and say that for all other values of x the definition fails. Or we may, as in Chap. I. § 8, leave values of x other than positive integral values entirely out of account, and regard our function as a function of the *positive integral variable* n, whose values are the positive integers

$$1, 2, 3, 4, \ldots.$$

In this case we may write

$$y = \phi(n)$$

and regard y now as a function of n defined for *all* values of n.

It is obvious that any function of x defined for all values of x gives rise to a function of n defined for all values of n. Thus from the function $y = x^2$ we deduce the function $y = n^2$ by merely omitting from consideration all values of x other than positive integers, and the corresponding values of y. On the other hand from any function of n we can deduce any number of functions of x by merely assigning values to y, corresponding to values of x other than positive integral values, in any way we please.

44. Interpolation. The problem of determining a function of x which shall assume, for all positive integral values of x, values agreeing with those of a given function of n, is of extreme importance in higher mathematics. It is called the *problem of functional interpolation.*

Were the problem however merely that of finding *some* function of x to fulfil the condition stated it would of course present no difficulty whatever. We could, as explained above, simply fill in the missing values as we pleased: we might indeed simply regard the given values of the function of n as *all* the values of the function of x and say that the definition of the latter function failed for all other values of x. But such purely theoretical solutions are obviously not what is usually wanted. What is usually wanted is some *formula* involving x (of as simple a kind as possible) which assumes the given values for $x = 1, 2, \ldots$.

In some cases, especially when the function of n is itself defined by a formula, there is an obvious solution. If for example $y = \phi(n)$, where $\phi(n)$ is a function of n which would have a meaning even were n not a positive integer (e.g. n, n^2, $(n-1)/(n+1)$), we naturally take our function of x to be $y = \phi(x)$. But even in this very simple case it is easy to write down other almost equally obvious solutions of the problem. For example

$$y = \phi(x) + \sin x\pi,$$

assumes the value $\phi(n)$ for $x = n$, since $\sin n\pi = 0$.

In other cases $\phi(n)$ may be defined by a formula, such as $(-1)^n$, which ceases to define for some values of x at any rate (as here in the case of fractional values of x with even denominators, or irrational values). But it may be possible to transform the formula in such a way that it does define for all values of x. In this case, for example,

$$(-1)^n = \cos n\pi,$$

if n is an integer, and the problem of interpolation is solved by the function $\cos x\pi$.

In other cases $\phi(x)$ may be defined for some values of x other than positive integers, but not for all. Thus from $y=n^n$ we are led to $y=x^x$. This expression has a meaning for some only of the remaining values of x. If for simplicity we confine ourselves to *positive* values of x, x^x has a meaning for all *rational* values of x, since

$$(p/q)^{p/q} = \sqrt[q]{(p/q)^p},$$

according to the definition of fractional indices adopted in elementary algebra. But when x is *irrational* x^x has (so far as we are in a position to say at the present moment) no meaning at all. Thus in this case the problem of interpolation at once leads us to consider the question of extending our definitions in such a way that x^x shall have a meaning even when x is irrational. We shall see later on how the desired extension may be effected.

Again consider the case in which

$$y = 1 \cdot 2 \ldots n = n!$$

In this case there is no obvious formula in x which reduces to $n!$ for $x=n$, as $x!$ means nothing for values of x other than the positive integers. This is a case in which attempts to solve the problem of interpolation have led to important advances in mathematics. For mathematicians have succeeded in discovering a function (the Gamma-function) which possesses the desired property and many other interesting and important properties besides.

45. Finite and infinite classes. Before we proceed further it is necessary to make a few remarks about certain ideas of an abstract and logical nature which are of constant occurrence in Pure Mathematics.

In the first place, the reader is probably familiar with the notion of **a class.** It is unnecessary to discuss here any logical difficulties which may be involved in the notion of a 'class': roughly speaking we may say that a class is the aggregate or collection of all the entities or objects which possess a certain property, simple or complex. Thus we have the class of British subjects, or red-headed Germans, or positive integers, or real numbers.

Moreover, the reader has probably an idea of what is meant by a **finite** or **infinite** class. Thus the class of *British subjects* is a finite class: the aggregate of all British subjects, past, present, and future, has a certain definite number n, though of course we cannot tell at present the actual value of n. The class of *present British subjects*, on the other hand, has a number n which could be ascertained by counting, were the methods of the census effective enough.

On the other hand the class of positive integers is *not* finite but infinite. This may be expressed more precisely as follows. If n is any positive integer (e.g. 1000, 1,000,000 or any number we like to think of), there are more than n positive integers. Thus if the number we think of is 1,000,000, there are obviously at least 1,000,001 positive integers. Similarly the class of real numbers, or of points on a line, is infinite. It is convenient to express this by saying that there are **an infinite number** of positive integers, or real numbers, or points on a line. But the reader must be careful always to remember that by saying this we mean **simply** that the class in question is not a class with a definite number of members, such as 1000 or 1,000,000.

46. Properties possessed by a function of n for large values of n. We may now return to the 'functions of n' which we were discussing in §§ 43, 44. They have many points of difference from the functions of x which we discussed in Chap. II. But there is one fundamental characteristic which the two classes of functions have in common—*the values of the variable for which they are defined form an infinite class.* It is this fact which forms the basis of all the considerations which follow and which, as we shall see in the next chapter, apply, *mutatis mutandis*, to functions of x as well.

Suppose that $\phi(n)$ is any function of n, and that P is any property which $\phi(n)$ may or may not have, such as that of being a positive integer or of being greater than 1. Consider, for each of the values $n = 1, 2, 3, \ldots$, whether $\phi(n)$ has the property P or not. Then there are three possibilities—

(*a*) $\phi(n)$ may have the property P for *all* values of n, or for all values of n except a definite number N of such values:

(*b*) $\phi(n)$ may have the property for *no* values of n, or *only* for a definite number N of such values:

(*c*) neither (*a*) nor (*b*) may be true.

If (*b*) is true the values of n for which $\phi(n)$ has the property form a finite class. If (*a*) is true the values of n for which $\phi(n)$ has *not* the property form a finite class. In the third case neither class is finite. Let us consider some particular cases.

(1) Let $\phi(n)=n$, and let P be the property of being a positive integer. Then $\phi(n)$ has the property P for *all* values of n.

If on the other hand P denotes the property of being a positive integer greater than or equal to 1000, $\phi(n)$ has the property for all values of n except a definite number of values of n, viz. 1, 2, 3, ..., 999. In either of these cases (*a*) is true.

(2) If $\phi(n)=n$, and P is the property of being less than 1000, (*b*) is true.

(3) If $\phi(n)=n$, and P is the property of being *odd*, (*c*) is true. For $\phi(n)$ is odd if n is odd and even if n is even, and either the odd or the even values of n form an infinite class.

Example. Consider, in each of the following cases, whether (*a*), (*b*), or (*c*) is true:—

(i) $\phi(n)=n$, P being the property of being a perfect square,

(ii) $\phi(n)=$ the nth prime number, P being the property of being odd,

(iii) $\phi(n)=$ the nth prime number, P being the property of being even,

(iv) $\phi(n)=$ the nth prime number, P being the property $\phi(n)>n$,

(v) $\phi(n)=1-(-1)^n(1/n)$, P being the property $\phi(n)<1$,

(vi) $\phi(n)=1-(-1)^n(1/n)$, P being the property $\phi(n)<2$,

(vii) $\phi(n)=1000\{1+(-1)^n\}/n$, P being the property $\phi(n)<1$,

(viii) $\phi(n)=1/n$, P being the property $\phi(n)<\cdot 001$,

(ix) $\phi(n)=(-1)^n/n$, P being the property $|\phi(n)|<\cdot 001$,

(x) $\phi(n)=10000/n$, or $(-1)^n 10000/n$, P being either of the properties $\phi(n)<\cdot 001$ or $|\phi(n)|<\cdot 001$,

(xi) $\phi(n)=(n-1)/(n+1)$, P being the property $1-\phi(n)<\cdot 0001$.

47. Let us now suppose that $\phi(n)$ and P are such that the assertion (*a*) is true, i.e. that $\phi(n)$ has the property P, if not for all values of n, at any rate for all values of n except a definite number N of such values. We may denote these exceptional values by

$$n_1,\ n_2,\ \dots,\ n_N.$$

There is of course no reason why these N values should be the

first N values $1, 2, \dots, N$, though, as the preceding examples show, this is frequently the case in practice.

But whether this is so or not we know that $\phi(n)$ has the property P if $n > n_N$. Thus the nth prime is odd if $n > 2$ ($n = 2$ being the only exception to the statement), and $1/n < \cdot 001$ if $n > 1000$ (the first 1000 values of n being the exceptions), and

$$1000\{1 + (-1)^n\}/n < 1,$$

if $n > 2000$, the exceptional values being $2, 4, 6, \dots, 2000$. That is to say, in each of these cases the property is possessed *for all values of* n *from a definite value onwards.*

We shall frequently express this by saying that $\phi(n)$ has the property for **large**, or *very large*, or *all sufficiently large* values of n. Thus when we say that $\phi(n)$ *has the property* P (which will as a rule be a property expressed by some relation of inequality) *for large values of* n, what we mean is that we can determine some definite number, n_0 say, such that $\phi(n)$ has the property for all values of n greater than or equal to n_0. This number n_0, in the examples considered above, may be taken to be any number greater than n_N, the greatest of the exceptional numbers: it is most natural to take it to be $n_N + 1$.

Thus we may say that 'all large primes are odd,' or that '$1/n$ is less than $\cdot 001$ for large values of n.' And the reader must make himself familiar with the use of the word *large* in statements of this kind. *Large* is in fact a word which, standing by itself, has no more absolute meaning in mathematics than in the language of common sense. It is a truism that in common life a number which is large in one connection is small in another; 6 goals is a large score in a football match, but 6 runs is not a large score in a cricket match; and 300 runs is a large score, but £300 is not a large income—and so of course in mathematics *large* generally means *large enough*, and what is large enough for one purpose may not be large enough for another.

We know now what is meant by the assertion '$\phi(n)$ has the property P for large values of n.' It is with assertions of this kind that we shall be concerned throughout this chapter. Given a function $\phi(n)$, are there any properties of which such an assertion is true?

48. The phrase 'n tends to infinity.' There is a somewhat different way of looking at the matter which it is natural to adopt. Suppose that n assumes successively the values $1, 2, 3, \dots$. The word 'successively' naturally suggests succession *in time*, and we may suppose n, if we like, to assume these values at successive moments of time (e.g. at the beginnings of successive seconds). Then as the seconds pass n gets larger and larger and there is no limit to the extent of its increase. However large a number we may think of (e.g. 969372855) a time will come when n has become larger than this number.

It is convenient to have a short phrase to express this unending growth of n, and we shall say that n **tends to infinity**, or $n \to \infty$, this last symbol being usually employed as an abbreviation for 'infinity.' The phrase 'tends to' like the word 'successively' naturally suggests the idea of change *in time*, and it is convenient to think of the variation of n as accomplished in time in the manner described above. This however is a mere matter of convenience. The variable n is a purely logical entity which has in itself nothing to do with time.

The reader cannot too strongly impress upon himself that when we say that n 'tends to ∞' we mean *simply* that n is supposed to assume a series of values which increase continually and without limit. **There is no number 'infinity':** such an equation as

$$n = \infty$$

is as it stands *absolutely meaningless*: n cannot be equal to ∞, because 'equal to ∞' means nothing. So far in fact the symbol ∞ means nothing at all except in the one phrase 'tends to ∞,' the meaning of which we have explained above. Later on we shall learn how to attach a meaning to other phrases involving the symbol ∞, but the reader will always have to bear in mind

(1) that ∞ *by itself* means nothing, although *phrases containing it* sometimes mean something,

(2) that in every case in which a phrase containing the symbol ∞ means something it will do so simply because we have previously attached a meaning to it by means of a special definition.

Now it is clear that if $\phi(n)$ has the property P for large values of n, and if n 'tends to ∞,' in the sense which we have just explained, n will ultimately assume values large enough to ensure that $\phi(n)$ has the property P. And so another way of putting the question 'what properties has $\phi(n)$ for sufficiently large values of n?' is 'how does $\phi(n)$ behave as n tends to ∞?'

49. The behaviour of a function of n as n tends to infinity. We shall now proceed, in the light of the remarks made in the preceding sections, to consider the meaning of some kinds of statements which are perpetually occurring in higher mathematics. Let us consider for example, the two following statements—(*a*) *$1/n$ is small for large values of n*, (*b*) *$1-(1/n)$ is nearly equal to 1 for large values of n*,—neither of which, we imagine, anyone will be inclined to dispute. Yet, obvious as they may seem, there is a good deal in them which will repay the reader's attention. Let us take (*a*) first, as being slightly the simpler.

We have already considered the statement '*$1/n$ is less than* ·01 *for large values of n*.' This, we saw, means that the inequality $1/n < \cdot 01$ is true for all values of n greater than some definite value, in fact greater than 100. Similarly it is true that '*$1/n$ is less than* ·0001 *for large values of n*': in fact $1/n < \cdot 0001$ if $n > 10000$. And instead of ·01 or ·0001 we might take ·000001 or ·00000001, or indeed any positive number we like.

It is obviously convenient to have some way of expressing the fact that *any* such statement as '*$1/n$ is less than* ·01 *for large values of n*' is true, when we substitute for ·01 some smaller number, such as ·0001 or ·000001 or any other number we care to choose. And clearly we can do this by saying that '*however small δ may be* (provided of course it is positive) *$1/n < \delta$ for sufficiently large values of n*.' That this is true is obvious. For $1/n < \delta$ if $n > 1/\delta$; so that our 'sufficiently large' values of n need only all be greater than $1/\delta$. The assertion is however a complex one, in that it really stands for the whole class of assertions which we obtain by giving to δ special values such as ·01. And of course the smaller is δ and the larger $1/\delta$ the larger must the least of the 'sufficiently large' values of n be, values which are sufficiently large when δ has one value being inadequate when it has another.

The last statement italicised is what is really meant by the statement (a), that $1/n$ is small when n is large. Similarly for (b), which really means "*if* $\phi(n) = 1 - (1/n)$, *then the statement* '$1 - \phi(n) < \delta$ *for sufficiently large values of* n' *is true whatever positive value (such as* ·01 *or* ·0001) *we attribute to* δ." That the statement (b) is true is obvious from the fact that $1 - \phi(n) = 1/n$.

There is another way in which it is common to state the facts expressed by the assertions (a) and (b). This is suggested at once by § 48. Instead of saying '$1/n$ is small for large values of n' we say '$1/n$ tends to 0 as n tends to ∞.' Similarly we say that '$1-(1/n)$ tends to 1 as n tends to ∞': and these statements are to be regarded as precisely equivalent to (a) and (b). Thus the statements

'$1/n$ is small when n is large,'

'$1/n$ tends to 0 as n tends to ∞,'

are equivalent to one another and to the more formal statement

'if δ is any positive number, however small, $1/n < \delta$ for sufficiently large values of n,'

or to the still more formal statement

'if δ is any positive number, however small, we can find a number n_0 such that $1/n < \delta$ for all values of n greater than or equal to n_0.'

The reader should imagine himself confronted by an opponent who questions the truth of the statement. He would name a series of numbers growing smaller and smaller. He might begin with ·001. The reader would reply that $1/n < {\cdot}001$ as soon as $n > 1000$. The opponent would be bound to admit this, but would try again with some smaller number, such as ·0000001. The reader would reply that $1/n < {\cdot}0000001$ as soon as $n > 10000000$: and so on. In this simple case it is evident that the reader would always have the better of the argument.

We shall now introduce yet another way of expressing this property of the function $1/n$. We shall say that '*the* **limit** *of* $1/n$ *as* n *tends to* ∞ *is* 0,' a statement which we may express symbolically in the form

$$\lim_{n \to \infty} (1/n) = 0,$$

or simply $\lim (1/n) = 0$. We shall also sometimes use the notation

$$1/n \underset{(n \to \infty)}{\to} 0,$$

or simply $1/n \to 0$, which may be read '$1/n$ tends to 0 as n tends to ∞.' In the same way we shall write

$$\lim_{(n\to\infty)} \{1-(1/n)\} = 1, \quad \lim \{1-(1/n)\} = 1, \quad 1-(1/n) \underset{(n\to\infty)}{\to} 1,$$

or $1-(1/n) \to 1$.

50. Now let us consider a different example: let $\phi(n) = n^2$. Then '*n^2 is large when n is large.*' This statement is equivalent to the more formal statements

'if G is any positive number, however large, $n^2 > G$ for sufficiently large values of n,'

'we can find a number n_0 such that $n^2 > G$ for all values of n greater than or equal to n_0.'

And it is natural in this case to say that 'n^2 tends to ∞ as n tends to ∞,' or 'n^2 tends to ∞ with n' and to write

$$n^2 \underset{(n\to\infty)}{\to} \infty,$$

or simply $n^2 \to \infty$.

Finally consider the function $\phi(n) = -n^2$. In this case $\phi(n)$ is large, but negative, when n is large, and we naturally say that '$-n^2$ tends to $-\infty$ as n tends to ∞' and write

$$-n^2 \to -\infty.$$

And the use of the symbol $-\infty$ in this sense suggests that it will sometimes be convenient to write $n^2 \to +\infty$ for $n^2 \to \infty$ and generally to use $+\infty$ instead of ∞, in order to secure greater uniformity of notation.

But we must once more repeat that in all these statements the symbols ∞, $+\infty$, $-\infty$ mean nothing whatever in themselves, and only acquire a meaning when they occur in certain special connections in virtue of the explanations which we have just given.

51. Definition of a limit. After the discussion which precedes the reader should be in a position to appreciate the general notion of a *limit*. Roughly we may say that *$\phi(n)$ tends to a limit l as n tends to ∞ if $\phi(n)$ is nearly equal to l when n is large.* But although the meaning of this statement should be clear enough after the preceding explanations, it is not, as it stands, precise enough to serve as a strict mathematical definition.

It is, in fact, equivalent to a whole class of statements of the type '*for sufficiently large values of* n, $\phi(n)$ *differs from* l *by less than* δ.' This statement has to be true for $\delta = \cdot 01$ or $\cdot 0001$ or *any* positive number; and for any such value of δ it has to be true for *any* value of n after a certain definite value n_0, though, the smaller δ, the larger (as a rule) will be this value n_0.

We accordingly frame the following formal definition:

DEFINITION I. *The function* $\phi(n)$ *is said to tend to the limit* l *as* n *tends to* ∞, *if, however small be the positive number* δ, $\phi(n)$ *differs from* l *by less than* δ *for sufficiently large values of* n; *or if, however small be the positive number* δ, *we can determine a value* n_0 *corresponding to* δ *and such that* $\phi(n)$ *differs from* l *by less than* δ *for all values of* n *greater than or equal to* n_0.

It is usual to denote the difference $\phi(n) \sim l$, taken positively, by $|\phi(n) - l|$. It is equal to $\phi(n) - l$ or to $l - \phi(n)$, whichever is positive, and agrees with the definition of the *modulus* of $\phi(n) - l$, as given in Ch. III, though at present we are only considering real values, positive or negative.

With this notation the definition may be stated more shortly as follows: '*if, given any positive number,* δ, *however small, we can find* n_0 *so that* $|\phi(n) - l| < \delta$ *for* $n \geqq n_0$, *then we say that* $\phi(n)$ *tends to the limit* l *as* n *tends to* ∞, *and write*

$$\lim_{n \to \infty} \phi(n) = l.\text{'}$$

Sometimes we may omit the '$n \to \infty$'; and sometimes it is convenient, for brevity, to write $\phi(n) \to l$.

It should be observed that n_0 is a function of δ. Thus if $\phi(x) = 1/n$, $l = 0$, and the condition reduces to $1/n < \delta$ $(n \geqq n_0)$, which is satisfied if $n_0 = 1 + [1/\delta]$ (the integer larger by one than the greatest integer contained in $1/\delta$). There is one and only one case in which *the same* n_0 will do for *all* values of δ. If, from a certain value N of n onwards $\phi(n)$ is constant, say equal to C, it is evident that $\phi(n) - C = 0$ for $n \geqq N$, so that the inequality $|\phi(n) - C| < \delta$ is satisfied for $n \geqq N$ and *all* positive values of δ. And if $|\phi(n) - l| < \delta$ for $n \geqq N$ and *all* positive values of δ it is evident that $\phi(n) = l$ for $n \geqq N$, so that $\phi(n)$ is constant for all such values of n.

52. The definition of a limit may be illustrated geometrically as follows. The graph of $\phi(n)$ consists of a number of points corresponding to the values $n = 1, 2, 3, \ldots$.

Draw the line $y = l$, and the parallel lines $y = l - \delta$, $y = l + \delta$ at distance δ from it. Then

$$\lim_{n \to \infty} \phi(n) = l,$$

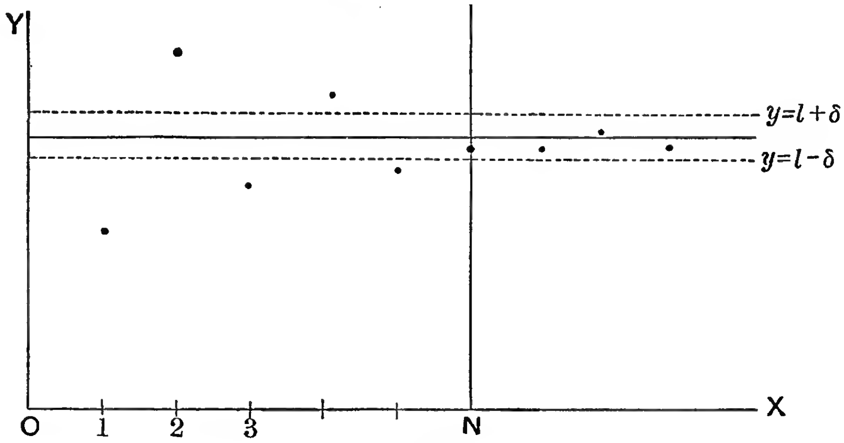


FIG. 35.

if, when once these lines have been drawn, no matter how close they may be together, we can always draw a line $x = N$ (as in the figure) in such a way that the point of the graph on this line, and all points to the right of it, lie between them. We shall find this geometrical way of looking at our definition particularly useful when we come to deal with functions defined for all values of a real variable and not merely for positive integral values.

53. So much for functions of n which tend to a limit as n tends to ∞. We must now frame corresponding definitions for functions which, like the functions n^2 or $-n^2$, tend to positive or negative infinity. The reader should by now find no difficulty in appreciating the point of

DEFINITION II. *The function $\phi(n)$ is said to tend to $+\infty$ (positive infinity) with n, if, when any number G, however large, is assigned, we can determine n_0 so that $\phi(n) > G$ for $n \geqq n_0$; or if, however large G may be, $\phi(n) > G$ for sufficiently large values of n.*

Another, less precise, form of statement is '*if we can make $\phi(n)$ as large as we please by sufficiently increasing n.*' This is open to the objection that it obscures a fundamental point, viz. that $\phi(n)$ must be greater than G for *all* values of $n \geqq n_0$, and not merely for *some* such values. But there is no harm in using this form of expression if we are clear what it means.

When $\phi(n)$ tends to $+\infty$ we write

$$\phi(n) \underset{(n\to\infty)}{\to} +\infty.$$

We may leave it to the reader to frame the corresponding definition for functions which tend to negative infinity.

54. Some points concerning the definitions. The reader should carefully observe the following points.

(1) We may obviously alter the values of $\phi(n)$ for any finite number of values of n, in any way we please, without in the least affecting the behaviour of $\phi(n)$ as n tends to ∞. For example $1/n$ tends to 0 as n tends to ∞. We may deduce any number of new functions from $1/n$ by altering a finite number of its values. For instance we may consider the function $\phi(n)$ which is equal to 3 for $n = 1, 2, 7, 9, 106, 107, 108, 237$ and equal to $1/n$ for all other values of n. For this function, just as for the original function $1/n$, $\lim \phi(n) = 0$. Similarly, for the function $\phi(n)$ which is equal to 3 if $n = 1, 2, 7, 9, 106, 107, 108, 237$, and to n^2 otherwise, it is true that $\phi(n) \to +\infty$.

(2) On the other hand we cannot as a rule alter an *infinite* number of the values of $\phi(n)$ without fundamentally affecting its behaviour as n tends to ∞. If for example we altered the function $1/n$ by changing its value to 1 whenever n is a multiple of 100 it would no longer be true that $\lim \phi(n) = 0$. So long as a finite number of values only were affected we could always choose the number n_0 of the definition so as to be greater than the greatest of the values of n for which $\phi(n)$ was altered. In the examples above, for instance, we could always take $n_0 > 237$, and indeed we should be compelled to do so as soon as our imaginary opponent of § 49 had assigned a value of δ as small as 3 (in the first example) or a value of G as great as 3 (in the second). But now *however* large n_0 may be there will be greater values of n for which $\phi(n)$ has been altered.

(3) In applying the test of Definition I. it is of course absolutely essential that we should have $|\phi(n) - l| < \delta$ not merely for $n = n_0$ but for $n \geqq n_0$, i.e. *for* n_0 *and for all larger values of* n. In the last example, given δ we can obviously choose n_0 so that $\phi(n) < \delta$ for $n = n_0$: we have only to choose a sufficiently large value of n which is not a multiple of 100. But when n_0 is thus

chosen it is not true that $\phi(n) < \delta$ for $n \geqq n_0$: all the multiples of 100 which are greater than n_0 are exceptions to this statement.

(4) If $\phi(n)$ is always greater than l we can replace $|\phi(n) - l|$ by $\phi(n) - l$. Thus the test whether $1/n$ tends to the limit 0 as n tends to ∞ is simply whether $1/n < \delta$ for $n \geqq n_0$. If however $\phi(n) = (-1)^n/n$, l is again 0, but $\phi(n) - l$ is sometimes positive and sometimes negative. In such a case we must state the condition in the form $|\phi(n) - l| < \delta$, in this particular case in the form $|\phi(n)| < \delta$.

(5) *The limit l may itself be one of the actual values of* $\phi(n)$. Thus if $\phi(n) = 0$ for all values of n it is obvious that $\lim \phi(n) = 0$. Again if we had (in (2) and (3) above) altered the value of the function (when n is a multiple of 100) to 0 instead of to 1 we should have obtained a function $\phi(n) = 0$ (n a multiple of 100), $\phi(n) = 1/n$ (otherwise). The limit of this function as n tends to ∞ is still obviously 0. This limit is itself the value of the function for an infinite number of values of n, viz. all multiples of 100.

On the other hand *the limit itself need not (and in general will not) be the value of the function for any value of n.* This is sufficiently obvious in the case of $\phi(n) = 1/n$. The limit is 0; but the function is never zero for any value of n.

The reader cannot impress these facts too strongly on his mind. **A limit is not a value of the function:** it is something quite distinct from these values, though defined by its relations to them. The limits may possibly be equal to some of the values of the function—whether this be so or not has absolutely nothing to do with the notion of the limit: it is, so to say, a mere accident.

For the functions $\phi(n) = 0,\ 1,$

the limit is equal to *all* the values of $\phi(n)$: for

$$\phi(n) = 1/n, \quad (-1)^n/n, \quad 1 + (1/n), \quad 1 + \{(-1)^n/n\}$$

it is not equal to *any* value of $\phi(n)$: for

$$\phi(n) = (\sin \tfrac{1}{2}n\pi)/n, \quad 1 + \{(\sin \tfrac{1}{2}n\pi)/n\}$$

(whose limits as n tends to ∞ are easily seen to be 0 and 1, since $\sin \frac{1}{2}n\pi$ is never numerically greater than 1) the limit is equal to the value which $\phi(n)$ assumes for all even values of n, but the

values assumed for odd values of n are all different from the limit and from one another.

(6) A function may be always numerically very large when n is very large without tending either to $+\infty$ or to $-\infty$. A sufficient illustration of this is given by $\phi(n)=(-1)^n n$ or $(-1)^n n^2$. A function can only tend to $+\infty$ or to $-\infty$ if, after a certain value of n, it maintains a constant sign.

Examples XXVI. Consider the behaviour of the following functions of x as n tends to ∞:

1. $\phi(n)=n^k$, where k is a positive or negative integer or rational fraction. If k is positive n^k tends to $+\infty$ with n. If k is negative $\lim n^k=0$. If $k=0$, $n^k=1$ for all values of n (by the definition of n^0). Hence $\lim n^k=1$.

The reader will find it instructive, even in an almost obvious case like this, to write down a formal proof that the conditions of our definitions are satisfied. Take for instance the case of $k>0$. Let G be any assigned number, however large. We wish to choose n_0 so that $n^k>G$ for $n \geqq n_0$. We have in fact only to take for n_0 any number greater than $\sqrt[k]{G}$. If e.g. $k=4$, $n^4>10000$ for $n \geqq 11$, $n^4>100000000$ for $n \geqq 101$, and so on.

From a geometrical point of view the matter may be stated as follows. If $k>0$ the graph of $y=x^k$ is of the general form of A in Fig. 36; if $k<0$ of the form of B; if $k=0$ it is a line C parallel to the axis of x. At present we are only concerned with the series of points marked on these curves.

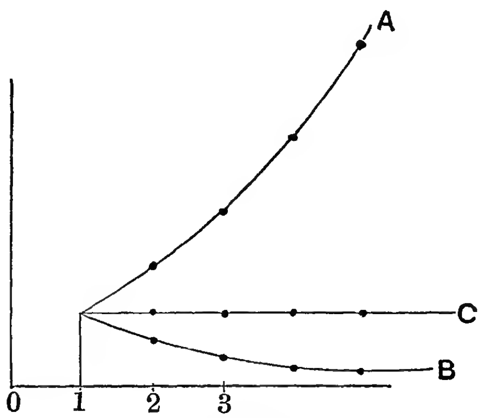


FIG. 36.

2. $\phi(n)=$ the nth prime number. If there were only a finite number of primes $\phi(n)$ would be defined only for a finite number of values of n. There are however, as was first shown by Euclid, infinitely many primes. Euclid's proof is as follows. If there are only a finite number of primes let them be $1, 2, 3, 5, 7, 11, \ldots N$. Consider the number $1+(1.2.3.5.7.11\ldots N)$. This number is evidently not divisible by any of $2, 3, 5, \ldots N$, since the remainder when it is divided by any of these numbers is 1. It is therefore

not divisible by any prime save 1, and is therefore itself prime, which is contrary to our hypothesis.

Again it is obvious that for all values of n (save $n=1$, 2, 3) $\phi(n)>n$. Hence $\phi(n)\to+\infty$.

3. $\phi(n)=$ the number of primes less than n. Here again $\phi(n)\to+\infty$.

4. $\phi(n)=[an]$, where a is any positive number. Here

$$\phi(n)=0 \quad (0\leqq n<1/a), \quad \phi(n)=1 \quad (1/a\leqq n<2/a),$$

and so on; and $\phi(n)\to+\infty$.

5. If $\phi(n)=1000000/n$, $\lim\phi(n)=0$: if $\psi(n)=n/1000000$, $\psi(n)\to+\infty$. These conclusions are in no way affected by the fact that at first $\phi(n)$ is much larger than $\psi(n)$ (being, in fact, larger until $n=1000$).

6. $\phi(n)=1/\{n-(-1)^n\}$, $n-(-1)^n$, $n\{1-(-1)^n\}$. The first function tends to 0, the second to $+\infty$, the third does not tend either to a limit or to $+\infty$.

7. $\phi(n)=(\sin n\theta)/n$, where θ is any real number. Since $|\sin n\theta|\leqq 1$, $|\phi(n)|<1/n$, and $\lim\phi(n)=0$.

8. $\phi(n)=(\sin n\theta)/\sqrt{n}$, $(\sin n\theta)/n^2$, $(a\cos^2 n\theta+b\sin^2 n\theta)/n$, where a and b are any real numbers.

9. $\phi(n)=\sin n\theta\pi$. If θ is integral $\phi(n)=0$ for all values of n, and therefore $\lim\phi(n)=0$.

Next let θ be rational, e.g. $\theta=p/q$, where p and q are positive integers. Let $n=aq+b$ where a is the quotient and b the remainder when n is divided by q. Then $\sin(np\pi/q)=(-1)^{ap}\sin(bp\pi/q)$. Suppose, for example, p even; then as n increases from 0 to $q-1$, $\phi(n)$ takes the values

$$0, \quad \sin(p\pi/q), \quad \sin(2p\pi/q), \ \dots \ \sin\{(q-1)p\pi/q\}.$$

When n increases from q to $2q-1$ these values are repeated; and so also as n goes from $2q$ to $3q-1$, $3q$ to $4q-1$, and so on. Thus the values of $\phi(n)$ form *a perpetual cyclic repetition of a finite series of different values.* It is evident that when this is the case $\phi(n)$ cannot tend to a limit or to $+\infty$ or to $-\infty$ as n tends to infinity.

The case in which θ is irrational is a little more difficult. It is discussed in the next set of examples.

55. Oscillating Functions. DEFINITION. *When $\phi(n)$ does not tend to a limit, nor to $+\infty$, nor to $-\infty$, as n tends to ∞, we say that $\phi(n)$* **oscillates** *as n tends to ∞.*

A function $\phi(n)$ certainly oscillates if its values form, as in the case considered in the last example above, a continual repetition of a cycle of values. But of course it may oscillate without possessing this peculiarity. Oscillation is, according to its definition, a **purely negative** quality—a function oscillates when it does not do certain other things.

The simplest example of an oscillatory function is given by

$$\phi(n) = (-1)^n,$$

which is equal to $+1$ when n is even and to -1 when n is odd. In this case the values recur cyclically. But consider

$$\phi(n) = (-1)^n + (1/n),$$

the values of which are

$$-1+1, \quad 1+(1/2), \quad -1+(1/3), \quad 1+(1/4), \quad -1+(1/5), \ldots$$

When n is large every value is nearly equal to $+1$ or -1, and obviously $\phi(n)$ does not tend to a limit or to $+\infty$ or to $-\infty$, and therefore it oscillates: but the values do not recur. It is to be observed that in this case every value of $\phi(n)$ is numerically less than or equal to $3/2$. Similarly

$$\phi(n) = (-1)^n 100 + (1000/n)$$

oscillates. When n is large enough every value is nearly equal to 100 or -100. The numerically greatest value is 900 (for $n=1$). But now consider $\phi(n) = (-1)^n n$, the values of which are -1, 2, -3, 4, -5, ... This function oscillates, for it does not tend to a limit, nor to $+\infty$, nor to $-\infty$. But in this case we cannot assign any limit beyond which the numerical value of the terms does not rise. The distinction between these two examples suggests a further definition.

DEFINITION. *If $\phi(n)$ oscillates as n tends to ∞ it will be said to* **oscillate finitely** *or* **infinitely** *according as it is or is not possible to assign a number K such that all the values of $\phi(n)$ are numerically less than K, i.e. $|\phi(n)| < K$ for all values of n.*

These definitions, as well as those of § 54, are further illustrated in the following examples.

Examples XXVII. Consider the behaviour as n tends to ∞ of the following functions:

1. $(-1)^n$, $5+3(-1)^n$, $(1000000/n)+(-1)^n$, $1000000(-1)^n+(1/n)$.

2. $(-1)^n n$, $1000000+(-1)^n n$. 3. $1000000-n$, $(-1)^n(1000000-n)$.

4. $n\{1+(-1)^n\}$. In this case the values of $\phi(n)$ are

$$0,\ 4,\ 0,\ 8,\ 0,\ 12,\ 0,\ 16,\ \ldots.$$

The odd terms are all zero and the even terms tend to $+\infty$: $\phi(n)$ oscillates infinitely.

5. $n^2+(-1)^n n$. The second term oscillates infinitely, but the first is very much larger than the second when n is large. In fact $\phi(n) \geqq n^2-n$ and $n^2-n=(n-\frac{1}{2})^2-\frac{1}{4}$ is greater than any assigned value G if $n>\frac{1}{2}+\sqrt{\{G+\frac{1}{4}\}}$. Thus $\phi(n)\rightarrow+\infty$.

6. $n^3+(-1)n^2$, $n^2\{1+(-1)^n\}$, $(-1)^n n^2+n$.

7. $11+3(-1)^n$, $(11/n)+3(-1)^n$, $11n+3(-1)^n$, $11+\{3(-1)^n/n\}$, $11+3n(-1)^n$, $\{11+3(-1)^n\}/n$, $\{11+3(-1)^n\}n$, $(11/n)+3(-1)^n n$, $11n+\{3(-1)^n/n\}$.

8. $\sin n\theta\pi$. We have already seen (Exs. XXVI. 9) that when θ is rational $\phi(n)$ oscillates finitely—unless θ is an integer, when $\phi(n)=0$, $\phi(n)\rightarrow 0$.

The case in which θ is irrational is a little more difficult. But it is not difficult to see that $\phi(n)$ still oscillates finitely. We can without loss of generality suppose $0<\theta<1$. In the first place $|\phi(n)|<1$. Hence $\phi(n)$ must oscillate finitely or tend to a limit. We shall consider whether the second alternative is really possible. Let us suppose then that

$$\lim \sin n\theta\pi=l.$$

Then however small ϵ is we can choose n_0 so that $\sin n\theta\pi$ lies between $l-\epsilon$ and $l+\epsilon$ for all values of n greater than or equal to n_0. Hence $\sin(n+1)\theta\pi-\sin n\theta\pi$ is numerically less than 2ϵ for all such values of n, and so $|\sin\frac{1}{2}\theta\pi\cos(n+\frac{1}{2})\theta\pi|<\epsilon$.

Hence $$\cos(n+\tfrac{1}{2})\theta\pi=\cos n\theta\pi\cos\tfrac{1}{2}\theta\pi-\sin n\theta\pi\sin\tfrac{1}{2}\theta\pi$$
must be numerically less than $\epsilon/|\sin\frac{1}{2}\theta\pi|$. Similarly
$$\cos(n-\tfrac{1}{2})\theta\pi=\cos n\theta\pi\cos\tfrac{1}{2}\theta\pi+\sin n\theta\pi\sin\tfrac{1}{2}\theta\pi$$
must be numerically less than $\epsilon/|\sin\frac{1}{2}\theta\pi|$; and so each of $\cos n\theta\pi\cos\frac{1}{2}\theta\pi$, $\sin n\theta\pi\sin\frac{1}{2}\theta\pi$ must be numerically less than $\epsilon/|\sin\frac{1}{2}\theta\pi|$. That is to say, if n is large $\cos n\theta\pi\cos\frac{1}{2}\theta\pi$ is very small, and this can only be the case if $\cos n\theta\pi$ is very small. Similarly $\sin n\theta\pi$ must be very small (so that l must be zero). But it is impossible that $\cos n\theta\pi$ and $\sin n\theta\pi$ can *both* be very small, as the sum of their squares is unity. Thus the hypothesis that $\sin n\theta\pi$ tends to a limit l is impossible, and therefore $\sin n\theta\pi$ oscillates as n tends to ∞.

The reader should consider with particular care the argument '$\cos n\theta\pi\cos\frac{1}{2}\theta\pi$ is very small, and this can only be the case if $\cos n\theta\pi$ is very small.' Why, he may ask, should it not be the other factor $\cos\frac{1}{2}\theta\pi$ which is 'very small'? The answer is to be found, of course, in the meaning of the phrase 'very small' as used in this connection. When we say '$\phi(n)$ is very small' for large values of n, we mean that we can choose n_0 so that $\phi(n)$ is numerically smaller than *any* assigned number, if n is sufficiently large. Such an assertion is palpably absurd when made of a *fixed* number such as $\cos\frac{1}{2}\theta\pi$, which is not zero.

9. $\sin n\theta\pi+(1/n)$, $\sin n\theta\pi+(1000000/n)$, $\sin n\theta\pi+1$, $\sin n\theta\pi+n$, $(-1)^n\sin n\theta\pi$, $\sin n\theta\pi+(-1)^n$, $\sin n\theta\pi+\{(-1)^n/n\}$, $\sin n\theta\pi+(-1)^n n$.

10. $a\cos n\theta\pi+b\sin n\theta\pi$, $\sin^2 n\theta\pi$, $\cos^2 n\theta\pi$, $a\cos^2 n\theta\pi+b\sin^2 n\theta\pi$.

11. $a+bn+(-1)^n(c+dn)+e\cos n\theta\pi+f\sin n\theta\pi$.

12. $n \sin n\theta\pi$. If n is integral, $\phi(n)=0$, $\phi(n)\to 0$. If n is rational but not integral, or irrational, $\phi(n)$ oscillates infinitely.

13. $n(a\cos^2 n\theta\pi + b\sin^2 n\theta\pi)$. In this case $\phi(n)\to+\infty$ if a and b are both positive, $\to-\infty$ if both are negative. Consider the special cases in which $a=0$, $b>0$, or $a>0$, $b=0$, or $a=0$, $b=0$. If a and b have opposite signs $\phi(n)$ generally oscillates infinitely. Consider any exceptional cases.

14. $\sin(n^2\theta\pi)$. If θ is integral, $\phi(n)\to 0$. Otherwise $\phi(n)$ oscillates finitely, as may be shown by arguments similar to though more complex than those used in Exs. XXVI. 9 and XXVII. 8*.

15. $\sin(n!\,\theta\pi)$. If θ has any rational value p/q, $n!\,\theta$ is certainly integral for all values of n greater than or equal to q. Hence $\phi(n)\to 0$. The case in which θ is irrational cannot be settled without the aid of considerations of a much more difficult character.

16. $\cos(n!\,\theta\pi)$, $a\cos^2(n!\,\theta\pi)+b\sin^2(n!\,\theta\pi)$, where θ is rational.

17. $an-[bn]$, $(-1)^n(an-[bn])$. 18. $[\sqrt{n}]$, $(-1)^n[\sqrt{n}]$, $\sqrt{n}-[\sqrt{n}]$.

19. *The smallest prime factor of n.* When n is a prime, $\phi(n)=n$. When n is even, $\phi(n)=2$. Thus $\phi(n)$ oscillates infinitely.

20. *The largest prime factor of n.*

21. *The number of days in the year n* A.D.

Examples XXVIII. 1. If $\phi(n)\to+\infty$ and $\psi(n)\geqq\phi(n)$ for all values of n, then $\psi(n)\to+\infty$.

2. If $\phi(n)\to 0$, and $|\psi(n)|\leqq|\phi(n)|$ for all values of n, then $\psi(n)\to 0$.

3. If $\lim|\phi(n)|=0$, then $\lim\phi(n)=0$.

4. If $\phi(n)$ tends to a limit or oscillates finitely, and $|\psi(n)|\leqq|\phi(n)|$ for $n\geqq n_0$, then $\psi(n)$ tends to a limit or oscillates finitely.

5. If $\phi(n)\to+\infty$, or $-\infty$, or oscillates infinitely, and $|\psi(n)|\geqq|\phi(n)|$ for $n\geqq n_0$, then $\psi(n)\to+\infty$ or $-\infty$ or oscillates infinitely.

6. 'If $\phi(n)$ oscillates and, however great be n_0 we can find values of n greater than n_0 and for which $\psi(n)$ is either greater than or less than $\phi(n)$, then $\psi(n)$ oscillates.' Is this true? If not give an example to the contrary [$\phi(n)=(-1)^n$, $\psi(n)=0$].

7. If $\phi(n)\to l$ as $n\to\infty$, then also $\phi(n+p)\to l$, p being any fixed integer. [This follows at once from the definition. Similarly we see that if $\phi(n)$ tends to $+\infty$ or $-\infty$ or oscillates so also does $\phi(n+p)$.]

8. The same conclusions hold (except in the case of oscillation) if p varies with n but is always numerically less than a fixed positive integer N; or if p varies with n in any way, so long as it is always positive.

9. Determine the least value of n_0 for which it is true that

(a) $n^2+n>1000$ $(n\geqq n_0)$, (b) $n^2+n>1000000$ $(n\geqq n_0)$.

* See Bromwich's *Infinite Series*, p. 485.

10. Determine the least value of n_0 for which it is true that

(a) $n+(-1)^n>1000 \quad (n \geqq n_0)$, (b) $n+(-1)^n>1000000 \quad (n \geqq n_0)$.

11. Determine the least value of n_0 for which it is true that

(a) $n^2+n>G \quad (n \geqq n_0)$, (b) $n+(-1)^n>G \quad (n \geqq n_0)$,

G being any positive number.

[(a) $n_0=[\frac{1}{2}+\sqrt{(G+\frac{1}{4})}]$: (b) $n_0=1+[G]$ or $2+[G]$, according as $[G]$ is odd or even: i.e. $n_0=1+[G]+\frac{1}{2}\{1+(-1)^{[G]}\}$.]

12. Determine the least value of n_0 such that

(a) $n/(n^2+1)<\cdot 0001$. (b) $(1/n)+\{(-1)^n/n^2\}<\cdot 000001$, for $n \geqq n_0$.

[Let us take the latter case. In the first place

$$(1/n)+\{(-1)^n/n^2\} \leqq (n+1)/n^2,$$

and it is easy to see that the least value of n_0 such that $(n+1)/n^2<\cdot 000001$, for $n \geqq n_0$, is 1000002. But the inequality given is satisfied by $n=1000001$, and this is the value of n_0 required.]

56. Some general theorems with regard to limits. A. The behaviour of the sum of two functions whose behaviour is known.

THEOREM I. *If $\phi(n)$ and $\psi(n)$ tend to limits a, b, then $\phi(n)+\psi(n)$ tends to the limit $a+b$.*

This is almost obvious. The argument which the reader will at once form in his mind is roughly this: 'when n is large $\phi(n)$ is nearly equal to a and $\psi(n)$ to b and therefore their sum is nearly equal to $a+b$.' It is well to state the argument quite formally, however.

Let δ be any assigned positive number (e.g. $\cdot 001, \cdot 0000001, \ldots$). We require to show that a number n_0 can be found such that

$$|\phi(n)+\psi(n)-a-b|<\delta \text{ ....(1)},$$

for $n \geqq n_0$. Now by a proposition proved in Chapter III. (more generally indeed than we need here) the modulus of the sum of two numbers is less than or equal to the sum of their moduli. Thus

$$|\phi(n)+\psi(n)-a-b| \leqq |\phi(n)-a|+|\psi(n)-b|.$$

It follows that the desired condition will certainly be satisfied if n_0 can be so chosen that

$$|\phi(n)-a|+|\psi(n)-b|<\delta \text{ ....(2)},$$

for $n \geqq n_0$. But this is certainly the case. For since $\lim \phi(n)=a$ we can, by the definition of a limit, find n_1 so that $|\phi(n)-a|<\delta'$,

for $n \geqq n_1$, and this however small may be δ'. Nothing prevents our taking $\delta' = \frac{1}{2}\delta$, so that $|\phi(n) - a| < \frac{1}{2}\delta$, for $n \geqq n_1$. Similarly we can find n_2 so that $|\psi(n) - b| < \frac{1}{2}\delta$ for $n \geqq n_2$. Now take n_0 to be *the greater of the two numbers* n_1, n_2. Then if $n \geqq n_0$ $|\phi(n) - a| < \frac{1}{2}\delta$ *and* $|\psi(n) - b| < \frac{1}{2}\delta$, and therefore (2) is satisfied and the theorem is proved.

The argument may be concisely stated thus: since $\lim \phi(n) = a$ and $\lim \psi(n) = b$ we can choose n_1, n_2 so that

$$|\phi(n) - a| < \tfrac{1}{2}\delta \ (n \geqq n_1), \quad |\psi(n) - b| < \tfrac{1}{2}\delta \ (n \geqq n_2),$$

and then, if n is greater than either n_1 or n_2,

$$|\phi(n) + \psi(n) - a - b| \leqq |\phi(n) - a| + |\phi(n) - b| < \delta,$$

and therefore $$\lim \{\phi(n) + \psi(n)\} = a + b.$$

Even when stated thus the argument may possibly appear to the reader to be merely a piece of useless pedantry, or an attempt to manufacture difficulties out of what is really obvious. We do not assert that such an opinion is, in this case, entirely groundless. The result really *is* very obvious: nor would any mathematician think it worth while as a rule to state arguments for what is so obvious at such length.

But the reader must remember that the theorem, obvious though it may be, is one of the most fundamental and important theorems in all mathematics. It is one which every mathematician uses, consciously or unconsciously, twenty times a day. The proof of such a theorem must be made absolutely clear, explicit, and rigorous: no room must be left for any possible misapprehension or confusion. And this is not all. The great majority of theorems concerning limits are, as the reader will discover before long, far from being so simple and so obvious as this one. In this case the result obviously indicated by common sense was true. In more difficult cases common sense as often indicates an untrue result as a true one: sometimes it fails to give any indication at all. In such cases vague general arguments are worse than useless: they lead to mistakes not only gross in themselves but entirely confusing in their consequences. And unless the reader is prepared to take the trouble to try and understand the way in which rigorous methods apply to simple and obvious cases, where their application is easy, he will find that when he comes to difficult questions, which cannot be settled without them, he has not the capacity to use them.

57. Results subsidiary to Theorem I. The reader should have no difficulty in verifying the following subsidiary results.

1. *If $\phi(n)$ tends to a limit, but $\psi(n)$ tends to $+\infty$ or to $-\infty$ or oscillates finitely or infinitely, then $\phi(n) + \psi(n)$ behaves like $\psi(n)$.*

2. *If* $\phi(n) \to +\infty$, *and* $\psi(n) \to +\infty$ *or oscillates finitely, then* $\phi(n)+\psi(n) \to +\infty$.

In this statement we may obviously change $+\infty$ into $-\infty$ throughout.

3. *But if* $\phi(n) \to \infty$ *and* $\psi(n) \to -\infty$, *then* $\phi(n)+\psi(n)$ *may either tend to a limit or to* $+\infty$ *or to* $-\infty$ *or may oscillate either finitely or infinitely.*

These five possibilities are illustrated in order by (i) $\phi(n)=n$, $\psi(n)=-n$, (ii) $\phi(n)=n^2$, $\psi(n)=-n$, (iii) $\phi(n)=n$, $\psi(n)=-n^2$, (iv) $\phi(n)=n+(-1)^n$, $\psi(n)=-n$, (v) $\phi(n)=n^2+(-1)^n n$, $\psi(n)=-n^2$. The reader should construct additional examples of each case.

4. *If* $\phi(n) \to +\infty$ *and* $\psi(n)$ *oscillates infinitely,* $\phi(n)+\psi(n)$ *may tend to* $+\infty$ *or oscillate infinitely, but cannot tend to a limit, or to* $-\infty$, *or oscillate finitely.*

For $\psi(n)=\{\phi(n)+\psi(n)\}-\phi(n)$; and, if $\phi(n)+\psi(n)$ behaved in any of the three last ways, it would follow, from the previous results, that $\psi(n) \to -\infty$, which is not the case. As examples of the two cases which are possible, consider (i) $\phi(n)=n^2$, $\psi(n)=(-1)^n n$, (ii) $\phi(n)=n$, $\psi(n)=(-1)^n n^2$. Here again the signs of $+\infty$ and $-\infty$ may be permuted throughout.

5. *If* $\phi(n)$ *and* $\psi(n)$ *both oscillate finitely,* $\phi(n)+\psi(n)$ *must tend to a limit or oscillate finitely.*

As examples take

(i) $\phi(n)=\psi(n)=(-1)^n$, (ii) $\phi(n)=\cos\frac{1}{3}n\pi$, $\psi(n)=\sin\frac{1}{3}n\pi$.

6. *If* $\phi(n)$ *oscillates finitely, and* $\psi(n)$ *infinitely, then* $\phi(n)+\psi(n)$ *oscillates infinitely.*

For $\phi(n)$ is in absolute value always less than a certain constant, say G. On the other hand $\psi(n)$, since it oscillates *infinitely* must assume values numerically greater than any assignable number (e.g. $10G$, $100G$, ...). Hence $\phi(n)+\psi(n)$ must assume values numerically greater than any assignable number (e.g. $9G$, $99G$, ...). Hence $\phi(n)+\psi(n)$ must either tend to $+\infty$ or $-\infty$ or oscillate infinitely. But if it tended to $+\infty$, for instance,

$$\psi(n)=\{\phi(n)+\psi(n)\}-\phi(n)$$

would also tend to $+\infty$, by the preceding results. Thus $\phi(n)+\psi(n)$ cannot tend to $+\infty$, nor, for similar reasons, to $-\infty$: hence it oscillates infinitely.

7. *If both* $\phi(n)$ *and* $\psi(n)$ *oscillate infinitely,* $\phi(n)+\psi(n)$ *may tend to a limit, or to* $+\infty$, *or to* $-\infty$, *or oscillate either finitely or infinitely.*

Suppose, for instance, that $\phi(n)=(-1)^n n$, while $\psi(n)$ is in turn each of the functions $(-1)^{n+1}n$, $\{1+(-1)^{n+1}\}n$, $-\{1+(-1)^n\}n$, $(-1)^{n+1}(n+1)$, $(-1)^n n$. We thus obtain examples of all five possibilities.

This exhausts all the possibilities which are really distinct. The results may be conveniently summarised in the following tabular form, in which 1 stands for 'tends to a limit,' 2 for 'tends to $+\infty$,' 3 for 'tends to $-\infty$,' 4 for 'oscillates finitely,' and 5 for 'oscillates infinitely.'

$\phi(n)$	$\psi(n)$	$\phi(n)+\psi(n)$
1	1	1
1	2	2
1	3	3
1	4	4
1	5	5
2	2	2
2	3	1, 2, 3, 4, or 5
2	4	2
2	5	2 or 5
4	4	1 or 4
4	5	5
5	5	1, 2, 3, 4, or 5

Before passing on to consider the *product* of two functions we may point out that the result of Theorem I. may be immediately extended to the sum of three or more functions which tend to limits as $n \to \infty$.

58. B. The behaviour of the product of two functions whose behaviour is known. We can now prove a similar set of theorems concerning the product of two functions. The principal result is the following.

THEOREM II. *If* $\lim \phi(n) = a$ *and* $\lim \psi(n) = b$, *then*

$$\lim \phi(n)\psi(n) = ab.$$

Let $$\phi(n) = a + \phi_1(n), \quad \psi(n) = b + \psi_1(n),$$

so that $\lim \phi_1(n) = 0$ and $\lim \psi_1(n) = 0$. Then

$$\phi(n)\psi(n) = ab + a\psi_1(n) + b\phi_1(n) + \phi_1(n)\psi_1(n).$$

Hence the numerical value of the difference $\phi(n)\psi(n) - ab$ is certainly not greater than the sum of the numerical values of $a\psi_1(n)$, $b\phi_1(n)$, $\phi_1(n)\psi_1(n)$. From this it is obvious that

$$\lim \{\phi(n)\psi(n) - ab\} = 0,$$

which proves the theorem.

The following is a strictly formal proof. We have

$$|\phi(n)\psi(n)-ab| \leqq |a\psi_1(n)|+|b\phi_1(n)|+|\phi_1(n)||\psi_1(n)|.$$

Choose n_0 so that for $n \geqq n_0$

$$|\phi_1(n)|<\tfrac{1}{3}\delta/|b|, \quad |\psi_1(n)|<\tfrac{1}{3}\delta/|a|.$$

Then $$|\phi(n)\psi(n)-ab|<\tfrac{1}{3}\delta+\tfrac{1}{3}\delta+\{\tfrac{1}{9}\delta^2/(|a||b|)\},$$

which is certainly less than δ, if $\delta<\frac{1}{3}|a||b|$. That is to say we can choose n_0 so that $|\phi(n)\psi(n)-ab|<\delta\ (n \geqq n_0)$, and so the theorem follows. The reader should study the details of this proof attentively; it is an elementary specimen of a type of proof perpetually occurring in higher analysis.

We need hardly point out that this theorem, like Theorem I., may be immediately extended to the product of any number of functions of n.

59. Results subsidiary to Theorem II. There is of course a series of theorems concerning products analogous to those stated in § 57 for sums. It will be convenient to present the results in tabular form. We must distinguish now *six* different ways in which $\phi(n)$ may behave as n tends to ∞. It may (1) tend to a limit *other than zero*, (2) tend to zero, (3) tend to $+\infty$, (3′) tend to $-\infty$, (4) oscillate finitely, (5) oscillate infinitely. We need not, as a rule, take account separately of (3) and (3′), as the results for one may be deduced from those for the other by a change of sign.

Case	$\phi(n)$	$\psi(n)$	$\phi(n)\psi(n)$
1	1	1	1
2	1	2	2
3	1	3	3 or 3′
4	1	4	4
5	1	5	5
6	2	2	2
7	2	3	any way
8	2	4	2
9	2	5	any way
10	3	3	3
11	3	4	3, 3′, or 5
12	3	5	3, 3′, or 5
13	4	4	1, 2, or 4
14	4	5	any way
15	5	5	any way

We leave the verification of this table as an exercise to the reader. The

more difficult cases 7, 9, 11, 13, 14, 15 may be illustrated by the following examples.

7

$\phi(n)$	$\psi(n)$	$\phi(n)\psi(n)$
$1/n$	n	1
$1/n^2$	n	2
$1/n$	n^2	3
$-1/n$	n^2	3′
$(-1)^n/n$	n	4
$(-1)^n/n$	n^2	5

9

$\phi(n)$	$\psi(n)$	$\phi(n)\psi(n)$
$(-1)^n/n$	$(-1)^n n$	1
$(-1)^n/n^2$	$(-1)^n n$	2
$(-1)^n/n$	$(-1)^n n^2$	3
$(-1)^n/n$	$(-1)^{n+1}n^2$	3′
$1/n$	$(-1)^n n$	4
$1/n$	$(-1)^n n^2$	5

11

$\phi(n)$	$\psi(n)$	$\phi(n)\psi(n)$
n	$2+(-1)^n$	3
n	$-2-(-1)^n$	3′
n	$(-1)^n$	5

13

$\phi(n)$	$\psi(n)$	$\phi(n)\psi(n)$
$(-1)^n$	$(-1)^n$	1
$1+(-1)^n$	$1-(-1)^n$	2
$\cos\frac{1}{3}n\pi$	$\sin\frac{1}{3}n\pi$	4

14

$\phi(n)$	$\psi(n)$	$\phi(n)\psi(n)$
$\{1-(-1)^n\}-[\{1+(-1)^n\}/n]$	$\{1-(-1)^n\}-n\{1+(-1)^n\}$	1
$1+(-1)^n$	$\{1-(-1)^n\}n$	2
$(-1)^n$	$(-1)^n n$	3
$(-1)^n$	$(-1)^{n+1}n$	3′
$1+(-1)^n$	$1+\{1-(-1)^n\}n$	4
$\cos\frac{1}{3}n\pi$	$n\sin\frac{1}{3}n\pi$	5

15

$\phi(n)$	$\psi(n)$	$\phi(n)\psi(n)$
$n\{1-(-1)^n\}-[\{1+(-1)^n\}/n]$	$[\{1-(-1)^n\}/n]-n\{1+(-1)^n\}$	1
$\{1+(-1)^n\}n$	$\{1-(-1)^n\}n$	2
$(-1)^n n$	$(-1)^n n$	3
$(-1)^n n$	$(-1)^{n+1}n$	3′
$\{1+(-1)^n\}n+[\{1-(-1)^n\}/n]$	$\{1-(-1)^n\}n$	4
$n\cos\frac{1}{3}n\pi$	$n\sin\frac{1}{3}n\pi$	5

As an illustration of how to verify these examples we may take the first example under Case 14. Since $1-(-1)^n=0$ or 2 according as n is even or odd, while $1+(-1)^n=0$ or 2 according as n is odd or even, the values of $\phi(n)$ are

$$2, \quad -2/2, \quad 2, \quad -2/4, \quad 2, \quad -2/6, \quad \ldots,$$

and so $\phi(n)$ oscillates finitely; while the values of $\psi(n)$ are

$$2, \quad -2\times 2, \quad 2, \quad -2\times 4, \quad 2, \quad -2\times 6, \quad \ldots,$$

and $\psi(n)$ oscillates infinitely. But $\phi(n)\psi(n)=4$ for all values of n.

60. A particular case of Theorem II which is important is that in which $\psi(n)$ is constant. The theorem then asserts simply that $\lim k\phi(n) = ka$ if $\lim \phi(n) = a$.

To this we may join the subsidiary theorem that if $\phi(n) \to +\infty$, then $a\phi(n) \to +\infty$ or $-\infty$, according as a is positive or negative, unless $a = 0$, when of course $a\phi(n) = 0$ for all values of n and $\lim a\phi(n) = 0$. And if $\phi(n)$ oscillates finitely or infinitely so does $a\phi(n)$, unless $a = 0$.

61. C. The behaviour of the difference or quotient of two functions whose behaviour is known. There is, of course, a similar set of theorems for the *difference* of two given functions: but they are such obvious corollaries from what precedes that it would be waste of time to state them at length. In order to deal with the quotient

$$\psi(n)/\phi(n),$$

we begin with the following theorem.

THEOREM III. *If* $\lim \phi(n) = a$, *and* a *is not zero, then*

$$\lim \{1/\phi(n)\} = 1/a.$$

Let
$$\phi(n) = a + \phi_1(n),$$
so that $\lim \phi_1(n) = 0$. Then

$$|\{1/\phi(n)\} - (1/a)| = |\phi_1(n)|/\{|a||a + \phi_1(n)|\},$$

and it is plain, since $\lim \phi_1(n) = 0$, that we can choose n_0 so that for $n \geqq n_0$ this is smaller than any assigned number δ.

The theorems subsidiary to this may again be stated concisely by means of a table.

Case	$\phi(n)$	$1/\phi(n)$
1	1	1
2	2	3, 3′, or 5
3	3	2
4	4	4 or 5
5	5	2, 4, or 5

The three more complicated cases may be illustrated by the following examples: the number indicates the behaviour of $1/\phi(n)$.

Case 2.	$\phi(n) = 1/n$	3
	$\phi(n) = -1/n$	3′
	$\phi(n) = (-1)^n/n$	5
Case 4.	$\phi(n) = (-1)^n$	4
	$\phi(n) = \{1+(-1)^n\} + [\{1-(-1)^n\}/n]$	5
Case 5.	$\phi(n) = (-1)^n n$	2
	$\phi(n) = \{1+(-1)^n\} + \{1-(-1)^n\} n$	4
	$\phi(n) = \{1+(-1)^n\} n + [\{1-(-1)^n\}/n]$	5

It will not be necessary now to attempt to state an exhaustive series of theorems for the quotient $\psi(n)/\phi(n)$, such as may be deduced from the results above and those of § 59. The principal theorem is

THEOREM IV. *If* $\lim \phi(n) = a$ *and* $\lim \psi(n) = b$, *and* $a \neq 0$, *then*

$$\lim \{\psi(n)/\phi(n)\} = b/a.$$

This requires no proof, being an immediate consequence of Theorems II and III.

The reader will however find it very instructive to draw up, at any rate partially, a table for the quotient similar to those we have given for the product and reciprocal, and to illustrate some of the possible cases with examples.

62. THEOREM V. *If* $R\{\phi(n), \psi(n), \chi(n), \dots\}$ *is any rational function of* $\phi(n)$, $\psi(n)$, $\chi(n)$, *etc., i.e. any function of the form*

$$P\{\phi(n), \psi(n), \chi(n), \dots\}/Q\{\phi(n), \psi(n), \chi(n), \dots\},$$

where P *and* Q *denote polynomials in* $\phi(n)$, $\psi(n)$, $\chi(n)$, ...: *and if*

$$\lim \phi(n) = a, \quad \lim \psi(n) = b, \quad \lim \chi(n) = c, \dots,$$

and

$$Q(a, b, c, \dots) \neq 0;$$

then

$$\lim R\{\phi(n), \psi(n), \chi(n), \dots\} = R(a, b, c, \dots).$$

For P is a sum of a finite number of terms of the type

$$A\{\phi(n)\}^p \{\psi(n)\}^q \dots,$$

where A is a constant and p, q positive integers. This term, by Theorem II (or rather by its obvious extension to the product of any number of functions) tends to the limit $Aa^p b^q \dots$, and so P tends to the limit $P(a, b, c, \dots)$, by the similar extension of Theorem I. Similarly for Q: and the result then follows from Theorem IV.

63. The preceding general theorem may be applied to the following very important particular problem: *what is the behaviour of the most general rational function of* n, viz.

$$S(n)=\frac{a_0n^p+a_1n^{p-1}+\dots+a_p}{b_0n^q+b_1n^{q-1}+\dots+b_q},$$

as n *tends to* ∞.

In order to apply the theorem we transform $S(n)$ by writing it in the form

$$n^{p-q}\left(a_0+\frac{a_1}{n}+\dots+\frac{a_p}{n^p}\right)\Big/\left(b_0+\frac{b_1}{n}+\dots+\frac{b_q}{n^q}\right).$$

The term in brackets is of the form $R\{\phi(n)\}$, where $\phi(n)=1/n$, and so, by Theorem V, it tends, as n tends to ∞, to the limit $R(0)=a_0/b_0$. Now, if $p<q$, $\lim n^{p-q}=0$; if $p=q$, $n^{p-q}=1$ and $\lim n^{p-q}=1$; if $p>q$, $n^{p-q}\to+\infty$. Hence, by Theorem II,

$$\lim S(n)=0\quad(p<q),\qquad \lim S(n)=a_0/b_0\quad(p=q),$$
$$S(n)\to+\infty\ \ (p>q,\ a_0/b_0\ \textit{positive}),\qquad S(n)\to-\infty\ \ (p>q,\ a_0/b_0\ \textit{negative}).$$

Examples XXIX. 1. Determine the behaviour, as $n\to\infty$, of each of the following functions of n, and of their sums, differences, products and quotients, taken in pairs: $1+\{(-1)^n/n\}$, $(-1)^n+(1/n)$, $1+(-1)^n n$, $(-1)^n+n$, $n+\{(-1)^n/n\}$, $(-1)^n n+(1/n)$, $(-1)^n(1+n)$, $(-1)^n\{1+(1/n)\}$, $(-1)^n\{n+(1/n)\}$.

2. Do the same for the functions

$$\cos^2\tfrac{1}{2}n\pi+(\sin^2\tfrac{1}{2}n\pi)/n,\quad \cos^2\tfrac{1}{2}n\pi+n\sin^2\tfrac{1}{2}n\pi,\quad n\cos^2\tfrac{1}{2}n\pi+(\sin^2\tfrac{1}{2}n\pi)/n.$$

3. Which (if any) of the functions

$$1/(\cos^2\tfrac{1}{2}n\pi+n\sin^2\tfrac{1}{2}n\pi),\quad 1/\{n(\cos^2\tfrac{1}{2}n\pi+n\sin^2\tfrac{1}{2}n\pi)\},$$
$$(n\cos^2\tfrac{1}{2}n\pi+\sin^2\tfrac{1}{2}n\pi)/\{n(\cos^2\tfrac{1}{2}n\pi+n\sin^2\tfrac{1}{2}n\pi)\}$$

tend to a limit as $n\to\infty$?

4. Denoting by $S(n)$ the general rational function of n, considered in § 63, show that in all cases

$$\lim\{S(n+1)/S(n)\}=1,\quad \lim[S\{n+(1/n)\}/S(n)]=1.$$

64. Functions of n which increase steadily with n. A special but particularly important class of functions of n is formed of those whose variation as n tends to ∞ is always in the same direction, that is to say those which always increase (or always decrease) as n increases. Since $-\phi(n)$ always increases if $\phi(n)$ always decreases, it is not necessary to consider the two kinds of functions separately; for theorems proved for one kind can at once be extended to the other.

DEFINITION. *The function $\phi(n)$ will be said to increase steadily with n if $\phi(n+1) \geqq \phi(n)$ for all values of n.*

It is to be observed that we do not exclude the case in which $\phi(n)$ has the *same* value for several values of n; all we exclude is possible *decrease*. Thus the function

$$\phi(n) = 2n + (-1)^n,$$

whose values for $n = 0, 1, 2, 3, 4, \dots$ are

$$1, 1, 5, 5, 9, 9, \dots$$

is said to increase steadily with n. Our definition would indeed include even functions which, from some value of n, remain constant; thus $\phi(n) = 1$ steadily increases according to our definition. However, as these functions are extremely special ones, and as there can be no doubt as to their behaviour as n tends to ∞, this apparent incongruity in the definition is not a serious defect.

There is one exceedingly important theorem concerning functions of this class.

THEOREM. *If $\phi(n)$ steadily increases with n, then either* (i) *$\phi(n)$ tends to a limit as n tends to ∞, or* (ii) *$\phi(n) \to +\infty$.*

That is to say, while there are in general *five* alternatives as to the behaviour of a function, there are *two* only for this special kind of function.

The proof is very simple. Imagine the various values of $\phi(n)$ represented by points along the line L of Chap. I. Each point lies to the *right* of the preceding point (or coincides with it).

Let P_n be the point corresponding to $\phi(n)$. Let Q be any other point whatever on the line. Then *either*

(1) there are values of n such that P_n lies to the right of Q (or coincides with it), *or*

(2) there are no such values.

In the first case we say that Q is a point which is *reached* for some value of n, in the second case that it is a point which is *not reached*. Every point is a reached point or a not reached point. If any point Q is reached, so obviously are all points to the left of Q.

There are two alternatives: (1) *every* point may be reached. Then it is clear that if G is any number, however large, it will correspond to a point Q, and, for sufficiently large values of n, P_n

will lie to the right of Q, and so will $P_{n+1}, \ldots$. In other words $\phi(n) > G$ *for all values of* n *from a certain value.* That is

$$\phi(n) \rightarrow +\infty.$$

Or (2) *not every* point may be reached. Then we can divide L into two segments, L_1, L_2, of which the first includes all reached points, the second all not reached points. The only doubt is as to whether the point R which divides the two segments is reached or not.

If R is reached, then, since no point P_n can lie to the right of R (as in that case other points in L_2 would be reached), all the points P_n must coincide with R from some value of n, the first for which R is reached. Thus if $OR = l$, we have $\phi(n) = l$ from a certain value of n onwards; so that, of course, $\lim \phi(n) = l$.

On the other hand, if R is not reached all points to the left of R, however close to R, are reached. Thus we can choose n_0 so that $\phi(n)$ is as nearly equal to l as we please when $n = n_0$. Since, as n increases beyond n_0, $\phi(n)$ approaches even more nearly to the value l, it is clear that

$$\lim \phi(n) = l.$$

The theorem is thus proved.

For example, if $\phi(n) = 3 - (1/n)$, $l = 3$: the point R ($OR = 3$) is not reached. From a common-sense point of view the theorem may be stated thus.

Let the point P move along the line L in a series of jumps, *its motion always being from left to right.* Then either P will pass over the whole line, or its position will gradually approximate to a definite position R on the line L. The theorem is almost intuitive: the proof which precedes is merely a careful analysis of the process of argument implied in but suppressed by our intuition of its truth.

Cor. 1. *If* $\phi(n)$ *increases steadily with* n *it will tend to a limit or to* $+\infty$ *according as it is or is not possible to find a fixed number* G *such that* $\phi(n) < G$.

We shall find this corollary exceedingly useful later on.

Cor. 2. *If* $\phi(n)$ *increases steadily with* n *and* $\phi(n) < G$ *for all values of* n, $\phi(n)$ *tends to a limit and this limit is less than or equal to* G.

It should be noticed that the limit may be equal to G: if e.g. $\phi(n) < 3 - (1/n)$, every value of $\phi(n)$ is less than 3, but the limit is equal to 3.

The reader should write out for himself the corresponding theorems and corollaries for the case in which $\phi(n)$ *decreases* as n increases.

65. The great importance of these theorems lies in the fact that they give us (what we have so far been without) a means of deciding (in a great many cases) whether a given function of n does or does not tend to a limit as $n \to \infty$, *without requiring us to be able to guess or otherwise infer beforehand what the limit is.* If we know what the limit must be (if there is one) we can use the test

$$|\phi(n) - l| < \epsilon \qquad (n \geqq n_0):$$

as for example in the case of $\phi(n) = 1/n$, where it is obvious that the limit can only be zero.

But suppose we have to determine whether

$$\phi(n) = \left(1 + \frac{1}{n}\right)^n$$

tends to a limit. In this case it is not obvious what the limit, if there is one, will be: and it is evident that the test above, which involves l, cannot, at any rate directly, be used to decide whether l exists or not.

Of course the test can sometimes be used indirectly, to prove that l *cannot* exist by means of a *reductio ad absurdum.* If e.g. $\phi(n) = (-1)^n$, it is clear that l would have to be equal to 1 and also equal to -1, which is obviously impossible.

66. The limit of x^n as n tends to ∞. Let us apply some of the preceding results to the particularly important case in which $\phi(n) = x^n$.

First, suppose x positive. Then since $\phi(n+1) = x\phi(n)$, $\phi(n)$ increases with n if $x > 1$, decreases as n increases if $x < 1$. If $x = 1$, $\phi(n) = 1$, $\lim \phi(n) = 1$, so that this special case need not detain us.

Thus, if $x > 1$, x^n must tend either to a limit (which must obviously be greater than 1), or to $+\infty$. Suppose it tends to a limit l. Then (Ex. XXVIII. 7) $\lim \phi(n+1) = \lim \phi(n) = l$; but

$$\lim \phi(n+1) = \lim x\phi(n) = x \lim \phi(n) = xl,$$

and therefore $l = xl$: and as x and l are both greater than 1, this is impossible. Hence

$$x^n \to +\infty \qquad (x > 1).$$

Ex. The reader may give an alternative proof, showing by the binomial theorem that, if $x = 1 + \delta$ $(\delta > 0)$, $x^n > 1 + n\delta$, and so that $x^n \to +\infty$.

On the other hand, if $x < 1$, x^n is a decreasing function and must therefore tend to a limit or to $-\infty$. Since x^n is positive the second alternative may be ignored. Thus $\lim x^n = l$, say, and as above $l = xl$, so that l must be zero. Hence

$$\lim x^n = 0 \qquad (0 < x < 1).$$

In the special cases of $x = 0$, 1, we clearly have $\lim x^n = 0$, $\lim x^n = 1$ respectively.

Ex. Prove as in the preceding example that, if $0 < x < 1$, $(1/x)^n$ tends to $+\infty$, and deduce that x^n tends to 0.

We have finally to consider the case in which x is negative. If $-1 < x < 0$ and $x = -y$, we have $\lim y^n = 0$ by what precedes and therefore $\lim x^n = 0$.

If $x = -1$ it is obvious that x^n oscillates, taking the values -1, 1 alternatively.

Finally if $x < -1$, $y > 1$, y^n tends to $+\infty$ and therefore x^n takes values, both positive and negative, numerically greater than any assigned number. Hence x^n oscillates infinitely.

To sum up:

$$\begin{array}{lr} \phi(n) = x^n \to +\infty, & (x > 1), \\ \lim \phi(n) = 1, & (x = 1), \\ \lim \phi(n) = 0, & (-1 < x < 1), \\ \phi(n) \text{ oscillates finitely}, & (x = -1), \\ \phi(n) \text{ oscillates infinitely}, & (x < -1). \end{array}$$

Examples XXX. 1. If $\phi(n)$ is positive and $\phi(n+1) > K\phi(n)$, where $K > 1$, for all values of n, then $\phi(n) \to +\infty$.

[For $$\phi(n) > K\phi(n-1) > K^2\phi(n-2) \ldots > K^{n-1}\phi(1),$$
from which the conclusion follows at once.]

2. The same result is true if the conditions above stated are satisfied only for $n \geqq n_0$.

3. If $\phi(n)$ is positive and $\phi(n+1) < K\phi(n)$, where $0 < K < 1$, then $\lim \phi(n) = 0$. This result also is true if the conditions are satisfied only for $n \geqq n_0$.

4. If $|\phi(n+1)|<K|\phi(n)|$ for $n\geqq n_0$, where $0<K<1$, then $\lim \phi(n)=0$.

5. If $\phi(n)$ is positive and $\lim\{\phi(n+1)\}/\{\phi(n)\}=l>1$, then $\phi(n)\rightarrow+\infty$.

[For we can determine n_0 so that $\{\phi(n+1)\}/\{\phi(n)\}>K>1$, for $n\geqq n_0$: we may, e.g., take K half-way between 1 and l. Now apply Ex. 1.]

6. If $\lim\{\phi(n+1)\}/\{\phi(n)\}=l$, where l is numerically less than unity, then $\lim\phi(n)=0$. [This follows from Ex. 4 as Ex. 5 follows from Ex. 1.]

7. Determine the behaviour, as $n\rightarrow\infty$, of $\phi(n)=n^r x^n$, where r is any positive integer.

[Here $$\{\phi(n+1)\}/\{\phi(n)\}=\{(n+1)/n\}^r x\rightarrow x$$ as $n\rightarrow\infty$. If x is positive and greater than 1, $\phi(n)\rightarrow+\infty$. If x is positive and $0<x<1$, $\phi(n)\rightarrow 0$. If x is negative and equal to $-y$, $\phi(n)=(-1)^n n^r y^n$, and it is easy to see that $\phi(n)\rightarrow 0$ $(-1<x<0)$ and $\phi(n)$ oscillates infinitely $(x\leqq -1)$. Finally if $x=1$, $\phi(n)=n^r$, and $\phi(n)\rightarrow+\infty$; and if $x=0$, $\phi(n)=0$ for all values of n.]

8. Discuss $n^{-r}x^n$ in the same way. [The results are the same, except that when $x=1$ or -1, $\phi(n)\rightarrow 0$.]

9. Draw up a table to show how $n^k x^n$ behaves as $n\rightarrow\infty$, for all real values of x, and all (positive and negative) integral values of k.

[The reader will observe that *the value of k is immaterial* except in the special cases when $x=1$ or -1. In other words, *it is the factor x^n which is the most important factor*: the second factor only asserts itself in the special cases when, owing to the fact that $x=\pm 1$, the first factor loses all or most of its importance. The fact is that since $\lim\{(n+1)/n\}^k=1$ for all values of k, positive or negative, the limit of the ratio $\phi(n+1)/\phi(n)$ depends only upon x.]

10. Prove that if x is positive $\sqrt[n]{x}\rightarrow 1$, as $n\rightarrow\infty$. [Suppose, e.g., $x>1$. Then $x, \sqrt{x}, \sqrt[3]{x}, \dots$ is a *decreasing* sequence, and $\sqrt[n]{x}>1$ for all values of n. Thus $\sqrt[n]{x}\rightarrow l$, where $l\geqq 1$. But if $l>1$ we could find values of n, as large as we please, for which $\sqrt[n]{x}>l$ or $x>l^n$: and as $l^n\rightarrow+\infty$ as $n\rightarrow\infty$ this is impossible.]

11. $\sqrt[n]{n}\rightarrow 1$. [For $\sqrt[n+1]{(n+1)}<\sqrt[n]{n}$ if $(n+1)^n<n^{n+1}$ or $\{1+(1/n)\}^n<n$, which is certainly satisfied if $n\geqq 3$ (see § 67 for a proof). Thus $\sqrt[n]{n}$ decreases as n increases from 3 onwards, and as it is always greater than unity it tends to a limit which is greater than or equal to unity. But if $\sqrt[n]{n}\rightarrow l\ (l>1)$, $n>l^n$, which is certainly untrue for sufficiently large values of n, since $l^n/n\rightarrow+\infty$ with n (Exs. 7, 8).]

12. $\sqrt[n]{(n!)}\rightarrow+\infty$. [However large G may be, $n!>G^n$ if n is large enough. For if $u_n=G^n/n!$, $u_{n+1}/u_n=G/n+1$, which tends to zero as $n\rightarrow\infty$, so that u_n does the same (Ex. 6).]

67. The limit of $\left(1+\dfrac{1}{n}\right)^n$. A more difficult case which can be settled by the help of § 64 is given by $\phi(n)=\{1+(1/n)\}^n$.

We shall prove first that

$$\left(1+\frac{1}{n}\right)^n > \left(1+\frac{1}{n-1}\right)^{n-1} \quad \text{..................(1)},$$

i.e. that

$$\left(\frac{n+1}{n}\right)^{n/(n-1)} - 1 > \frac{1}{n-1} \quad \text{..................(2)}.$$

Let

$$\left(\frac{n+1}{n}\right)^{1/(n-1)} = \alpha,$$

so that $\alpha > 1$. Then the inequality (2) may be written in the form

$$\alpha^n - 1 > \frac{n}{n-1}\left(\frac{n+1}{n} - 1\right),$$

or

$$(\alpha^n - 1)/n > (\alpha^{n-1} - 1)/(n-1) \quad \text{..............(3)},$$

or, dividing by the positive factor $\alpha - 1$,

$$(\alpha^{n-1} + \alpha^{n-2} + \dots + 1)/n > (\alpha^{n-2} + \alpha^{n-3} + \dots + 1)/(n-1) \quad \text{...(4)}.$$

Multiplying up and subtracting we obtain

$$(n-1)\,\alpha^{n-1} - \alpha^{n-2} - \alpha^{n-3} - \dots - 1 > 0 \quad \text{........(5)},$$

and this inequality is evidently true, since $\alpha > 1$. Thus the inequality (1) is established. Hence, by the theorem of § 65, $\{1 + (1/n)\}^n$ tends to a limit, or to $+\infty$, as $n \to \infty$.

But

$$\left(1+\frac{1}{n}\right)^n = 1 + n\,.\,\frac{1}{n} + \frac{n(n-1)}{1\,.\,2}\frac{1}{n^2} + \dots + \frac{n(n-1)\dots(n-n+1)}{1\,.\,2\dots n}\frac{1}{n^n},$$

by the binomial theorem; and so

$$\left(1+\frac{1}{n}\right)^n < 1 + 1 + \frac{1}{1\,.\,2} + \frac{1}{1\,.\,2\,.\,3} + \dots + \frac{1}{1\,.\,2\,.\,3\dots n}$$

$$< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}} < 3.$$

Thus $\{1 + (1/n)\}^n$ cannot tend to $+\infty$, and so

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e,$$

where e is a number such that $2 < e \leqq 3$. We shall have a great deal to do with this number e later on.

68. The limit of $n(\sqrt[n]{\beta} - 1)$. We proved above that if $\alpha > 1$

$$(\alpha^n - 1)/n > (\alpha^{n-1} - 1)/(n-1).$$

Let $\alpha^{n(n-1)} = \beta$. Then $\beta > 1$, and the inequality may be written in the form

$$(n-1)\left(\sqrt[n-1]{\beta} - 1\right) > n\left(\sqrt[n]{\beta} - 1\right).$$

Thus, if $\phi(n)=n(\sqrt[n]{\beta}-1)$, $\phi(n)$ *decreases* steadily as n increases. Also $\phi(n)$ is obviously always positive. Hence $\phi(n)$ tends to a limit l as $n\rightarrow\infty$, and $l\geqq 0$.

Moreover $\phi(n)>1-(1/\beta)$ for all values of n. For the inequality $n(\sqrt[n]{\beta}-1)>1-(1/\beta)$ becomes, if we put γ^n for β, $n\gamma^n(\gamma-1)>\gamma^n-1$, or

$$n\gamma^n>\gamma^{n-1}+\gamma^{n-2}+\dots+1,$$

which is obviously true, since $\gamma>1$. *Hence*

$$\lim_{n\rightarrow\infty} n(\sqrt[n]{\beta}-1)=f(\beta),$$

where $f(\beta)$ is a function of β, and $f(\beta)\geqq 1-(1/\beta)$ for all values of $\beta>1$.

Next suppose $\beta<1$, and let $\beta=1/\gamma$; then $n(\sqrt[n]{\beta}-1)=-n(\sqrt[n]{\gamma}-1)/\sqrt[n]{\gamma}$. As $n\rightarrow\infty$, $n(\sqrt[n]{\gamma}-1)\rightarrow f(\gamma)$, by what precedes. Also (Ex. XXX. 10)

$$\sqrt[n]{\gamma}\rightarrow 1.$$

Hence if $\beta=(1/\gamma)<1$,

$$n(\sqrt[n]{\beta}-1)\rightarrow -f(\gamma).$$

Finally, if $\beta=1$, $n(\sqrt[n]{\beta}-1)=0$ for all values of n.

Thus we arrive at the result: *the limit*

$$\lim n(\sqrt[n]{\beta}-1)$$

defines a function of β for all positive values of β. This function $f(\beta)$ possesses the properties

$$f(1/\beta)=-f(\beta),\quad f(1)=0,$$

and is positive or negative according as $\beta\gtrless 1$.

Later on we shall be able to identify this function as the *Napierian logarithm* of β.

Example. Prove that $f(a\beta)=f(a)+f(\beta)$. [Use the equations

$$f(a\beta)=\lim n(\sqrt[n]{a\beta}-1)=\lim\{n(\sqrt[n]{a}-1)\sqrt[n]{\beta}+n(\sqrt[n]{\beta}-1)\}.]$$

69. Infinite Series. Suppose that $u(n)$ is any function of n defined for all values of n. If we add up the values of $u(\nu)$ for $\nu=1, 2, \dots n$ we obtain another function of n, viz.

$$s(n)=u(1)+u(2)+\dots+u(n),$$

also defined for all values of n. It is generally most convenient to alter our notation slightly and write this equation in the form

$$s_n=u_1+u_2+\dots+u_n,$$

or, more shortly,

$$s_n=\sum_{\nu=1}^{n} u_\nu.$$

Now suppose that s_n tends to a limit s when n tends to ∞, i.e. that

$$\lim_{n\rightarrow\infty}\sum_{\nu=1}^{n} u_\nu=s.$$

This equation is usually written in one of the forms

$$\sum_{\nu=1}^{\infty} u_\nu = s, \quad u_1 + u_2 + u_3 + \dots = s,$$

the dots denoting the indefinite continuance of the series of u's.

The meaning of the above equations, expressed roughly, is that by adding more and more of the u's together we get nearer and nearer to the limit s. More precisely, if any small positive number ϵ is chosen, we can choose n_0 so that the sum of the first n_0 or any greater number of terms lies between $s-\epsilon$ and $s+\epsilon$; or in symbols

$$s - \epsilon < s_n < s + \epsilon,$$

if $n \geqq n_0$.

In these circumstances we shall call the series

$$u_1 + u_2 + \dots,$$

a **convergent infinite series,** and we shall call s the *sum* of the series, or the sum of *all* the terms of the series.

Thus to say that the series $u_0 + u_1 + \dots$ *converges and has the sum* s, or *converges to the sum* s or simply *converges to* s, is merely another way of stating that the sum $s_n = u_0 + u_1 + \dots + u_n$ of the first n terms tends to the limit s as $n \to \infty$, and the consideration of such infinite series introduces no new ideas beyond those with which the early part of this chapter should already have made the reader familiar. In fact the sum s_n is merely a function $\phi(n)$, such as we have been considering, expressed in a particular form. And any function $\phi(n)$ may be expressed in this form, by writing

$$\phi(n) = \phi(0) + [\phi(1) - \phi(0)] + \dots + [\phi(n) - \phi(n-1)].$$

It is sometimes convenient to say that $\phi(n)$ *converges to the limit* l, say, as $n \to \infty$. The use of the phrase 'converges' instead of 'tends to' is of course suggested by the phraseology usually employed in speaking of infinite series.

If $s_n \to +\infty$ or to $-\infty$ we shall say that the series $u_0 + u_1 + \dots$ is **divergent** or, *diverges to* $+\infty$, or $-\infty$, as the case may be. These phrases too may be applied to any function $\phi(n)$—e.g. if $\phi(n) \to +\infty$ we may say that $\phi(n)$ *diverges to* $+\infty$. If s_n does not tend to a limit or to $+\infty$ or to $-\infty$ it oscillates finitely or infinitely: in this case we say that the series $u_0 + u_1 + \dots$ oscillates finitely or infinitely.

70. General theorems concerning infinite series. When we are dealing with infinite series we shall constantly have occasion to use the following general theorems.

(1) If $u_1 + u_2 + \dots$ is convergent, and has the sum s, then $a + u_1 + u_2 + \dots$ is convergent and has the sum $a + s$. Similarly $a + b + c + \dots + k + u_1 + u_2 + \dots$ is convergent and has the sum $a + b + c + \dots + k + s$.

(2) If $u_1 + u_2 + \dots$ is convergent and has the sum s, then $u_{m+1} + u_{m+2} + \dots$ is convergent and has the sum

$$s - u_1 - u_2 - \dots - u_m.$$

(3) If any series considered in (1) or (2) diverges or oscillates so do the others.

(4) If $u_1 + u_2 + \dots$ is convergent and has the sum s, then $au_1 + au_2 + \dots$ is convergent and has the sum as.

(5) If the first series considered in (4) diverges or oscillates so does the second, unless $a = 0$.

(6) If $u_1 + u_2 + \dots$ and $v_1 + v_2 + \dots$ are both convergent the series $(u_1 + v_1) + (u_2 + v_2) + \dots$ is convergent and its sum is the sum of the first two series.

All these theorems are almost obvious and may be proved at once from the definitions or by applying the results of §§ 56–60 to the sum $s_n = u_1 + u_2 + \dots + u_n$.

(7) *If* $u_1 + u_2 + \dots$ *is convergent, then* $\lim u_n = 0$.

For $u_n = s_n - s_{n-1}$, and s_n and s_{n-1} have the same limit s. Hence $\lim u_n = s - s = 0$.

The reader may be tempted to think that the converse of the theorem is true and that if $\lim u_n = 0$ the series Σu_n must be convergent. That this is not the case is easily seen from an example. Let the series be

$$1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \dots$$

so that $u_n = 1/n$. The sum of the first four terms is

$$1 + \tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} > 1 + \tfrac{1}{2} + \tfrac{2}{4} = 1 + \tfrac{1}{2} + \tfrac{1}{2}.$$

The sum of the next four terms is $\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > \frac{4}{8} = \frac{1}{2}$; the sum of the next eight terms $> \frac{8}{16} = \frac{1}{2}$, and so on. The sum of the first

$$4 + 4 + 8 + 16 + \dots + 2^n = 2^{n+1}$$

terms is greater than

$$2 + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \dots + \tfrac{1}{2} = \tfrac{1}{2}(n+3),$$

and this increases beyond all limit with n: hence the series diverges to $+\infty$.

(8) If $u_1 + u_2 + u_3 + \dots$ is convergent, so is any series formed by grouping the terms in brackets in any way to form new single terms, as e.g. in $(u_1 + u_2 + u_3) + u_4 + (u_5 + u_6) + \dots$, and the sums of the two series are the same.

Here again the converse is not true. Thus, e.g. $1 - 1 + 1 - 1 + \dots$ oscillates, while $(1 - 1) + (1 - 1) + \dots$ or $0 + 0 + 0 + \dots$ converges to 0.

(9) *If every term u_n is positive* (*or zero*) *the series Σu_n must either converge or diverge to $+\infty$. If it converges its sum must be positive* (unless all the terms are zero, when of course its sum is zero).

For s_n is an increasing function of n, according to the definition of § 64, and we can apply the results of that section to s_n.

(10) *If every term u_n is positive* (*or zero*) *the necessary and sufficient condition that the series Σu_n should be convergent is that it should be possible to find a number G such that the sum of any number of terms is less than G, and if G can be so found the sum of the series is not greater than G.*

This also follows at once from § 64. It is perhaps hardly necessary to point out that the theorem is not true if the condition that every u_n is positive is not fulfilled. For example

$$1 - 1 + 1 - 1 + \dots$$

obviously oscillates, s_n being alternately equal to $+1$ and to 0.

(11) *If $u_1 + u_2 + \dots$, $v_1 + v_2 + \dots$ are two series of positive* (*or zero*) *terms, and the second series is convergent, and if $u_n \leqq v_n$ for all values of n, then the first series is also convergent, and its sum is less than or equal to that of the second.*

For, if $v_1 + v_2 + \dots = t$, $v_1 + v_2 + \dots + v_n \leqq t$, for all values of n, and so $u_1 + u_2 + \dots + u_n \leqq t$; which proves the theorem.

Conversely, if Σu_n is divergent, and $v_n \geqq u_n$, then Σv_n is divergent.

71. The infinite geometrical series. We shall now consider the 'geometrical' series, whose general term is $u_n = r^{n-1}$. In this case

$$s_n = 1 + r + r^2 + \dots + r^{n-1} = (1 - r^n)/(1 - r),$$

except in the special case in which $r = 1$, when

$$s_n = 1 + 1 + \dots + 1 = n.$$

In the last case $s_n \to +\infty$. In the general case s_n will tend to a

limit if and only if r^n does so. Referring to the results of § 66 we see that *the series* $1+r+r^2+\ldots$ *is convergent and has the sum* $1/(1-r)$ *if and only if* $-1<r<1$.

If $r \geqq 1$, $s_n \geqq n$, and so $s_n \rightarrow +\infty$; i.e. the series diverges to $+\infty$.

If $r=-1$, $s_n=1$ or 0 as n is odd or even: i.e. s_n oscillates finitely. If $r<-1$, s_n oscillates infinitely. Thus, to sum up, *the series* $1+r+r^2+\ldots$ *diverges* to $+\infty$ *if* $r \geqq 1$, *converges to* $1/(1-r)$ *if* $-1<r<1$, *oscillates finitely if* $r=-1$, *and oscillates infinitely if* $r<-1$.

Examples XXXI. 1. **Recurring decimals.** The commonest example of an infinite geometric series is given by an ordinary recurring decimal. Consider for example the decimal $\cdot 217\dot{1}\dot{3}$. This stands, according to the ordinary rules of arithmetic, for

$$\frac{2}{10}+\frac{1}{10^2}+\frac{7}{10^3}+\frac{1}{10^4}+\frac{3}{10^5}+\frac{1}{10^6}+\frac{3}{10^7}+\ldots=\frac{217}{1000}+\frac{13}{10^5}\bigg/\left(1-\frac{1}{10^2}\right)=\frac{2687}{12375}.$$

The reader should consider where and how any of the general theorems of § 70 have been used in this reduction.

2. Show that in general

$$\cdot a_1 a_2 \ldots a_m \dot{a}_1 a_2 \ldots \dot{a}_n = \frac{a_1 a_2 \ldots a_m a_1 \ldots a_n - a_1 a_2 \ldots a_n}{99\ldots 900\ldots 0},$$

the denominator containing n 9's and m 0's.

3. Show that a pure recurring decimal is always equal to a proper fraction whose denominator does not contain 2 or 5 as a factor.

4. A decimal with m non-recurring and n recurring decimal figures is equal to a proper fraction whose denominator is divisible by 2^m or 5^m but by no higher power of either. [For the decimal is converted into the sum of an integer and a pure recurring decimal by multiplication by 10^m, but not by multiplication by any lower power of 10.]

5. The converses of Exs. 3, 4 are also true, but their proof depends on Fermat's Theorem in the Theory of Numbers. If $r=p/q$, and q is prime to 10, it is known that we can find n so that 10^n-1 is divisible by q. Hence r may be expressed in the form $P/(10^n-1)$ or in the form

$$\frac{P}{10^n}+\frac{P}{10^{2n}}+\ldots$$

i.e. as a pure recurring decimal with n figures. But if $q=2^\alpha 5^\beta Q$, where Q is prime to 10, and m is the greater of α and β, $10^m r$ has a denominator prime to 10, and is therefore expressible as the sum of an integer and a pure recurring decimal. But this is not true of $10^\mu r$, for any value of μ less than m; hence the decimal for r has exactly m non-recurring figures.

6. To the results of Exs. 2—5 we must add that of Ex. I. 4. Finally, if we observe that

$$\cdot\dot{9}=\frac{9}{10}+\frac{9}{10^2}+\frac{9}{10^3}+\dots=1,$$

we see that every terminating decimal can also be expressed as a mixed recurring decimal whose recurring part is composed entirely of 9's. For example, $\cdot 217=\cdot 216\dot{9}$. Thus every proper fraction can be expressed as a recurring decimal, and conversely.

7. **Decimals in general. The expression of irrational numbers as non-recurring decimals.** Any decimal, whether recurring or not, corresponds to a definite number between 0 and 1. For the decimal $\cdot a_1 a_2 a_3 a_4 \dots$ stands for the series

$$\frac{a_1}{10}+\frac{a_2}{10^2}+\frac{a_3}{10^3}+\dots.$$

Since all the digits a_r are positive the sum s_n of the first n terms of this series increases with n: also it is certainly less than $\cdot\dot{9}$ or 1. Hence s_n tends to a limit between 0 and 1.

Moreover no two decimals can correspond to the same number (except in the special case noticed in Ex. 6). For suppose that $\cdot a_1 a_2 a_3 \dots$, $\cdot b_1 b_2 b_3 \dots$ are two decimals which agree as far as the figures a_{r-1}, b_{r-1}, while $a_r > b_r$. Then $a_r \geqq b_r + 1 > b_r . b_{r+1} b_{r+2} \dots$ (unless b_{r+1}, b_{r+2}, ... are all 9's), and so

$$\cdot a_1 a_2 \dots a_r a_{r+1} \dots > \cdot b_1 b_2 \dots b_r b_{r+1} \dots.$$

It follows that the expression of a rational fraction as a recurring decimal (Exs. 2—6) is unique. It also follows that every decimal which does *not* recur represents some *irrational* number between 0 and 1. Conversely, any such number can be expressed as such a decimal. For it must lie in one of the intervals

$$0,\ 1/10;\quad 1/10,\ 2/10;\quad \dots;\quad 9/10;\ 1.$$

If it lies in $r/10$, $(r+1)/10$ the first figure is r: by subdividing this interval into 10 parts we can determine the second figure; and so on.

Thus we see that the decimal $1\cdot 414\dots$, obtained by the ordinary process for the extraction of $\sqrt{2}$, cannot recur.

8. The decimals $\cdot 1010010001000010\dots$ and $\cdot 2020020002000020\dots$, in which the number of zeros between two 1's or 2's increases by one at each stage, represent irrational numbers.

9. The decimal $\cdot 11101010001010\dots$, in which the nth figure is 1 if n is prime, and zero otherwise, represents an irrational number. [Since the number of primes is infinite the decimal does not terminate. Nor can it recur: for if it did we could determine m and p so that m, $m+p$, $m+2p$, $m+3p$, ... are all prime numbers; and this is absurd, since the series includes $m+mp$.]*

* All the results of Exs. XXXI. may be extended, with suitable modifications, to decimals in any scale of notation. For a fuller discussion see Bromwich, *Infinite Series*, Appendix I.

Examples XXXII. 1. If $-1<r<1$, the series $r^m+r^{m+1}+\dots$ is convergent and its sum is $1/(1-r)-1-r-\dots-r^{m-1}$ (§ 70, (2)).

2. The series $r^m+r^{m+1}+\dots$ is convergent and its sum is $r^m/(1-r)$ (§ 70, (4)). Verify that the results of Exs. 1 and 2 are in agreement.

3. Prove that the series $1+2r+2r^2+\dots$ is convergent, and that its sum is $(1+r)/(1-r)$, (α) by writing it in the form $-1+2(1+r+r^2+\dots)$, (β) by writing it in the form $1+2(r+r^2+\dots)$, (γ) by adding the two series $1+r+r^2+\dots$, $r+r^2+\dots$. In each case mention which of the theorems of § 70 are used in your proof.

4. Prove that the arithmetic series

$$a+(a+b)+(a+2b)+\dots$$

is always divergent, unless both a and b are zero. Show that if $b\neq 0$ it diverges to $+\infty$ or to $-\infty$ according to the sign of b, while if $b=0$ it diverges to $+\infty$ or $-\infty$ according to the sign of a.

5. What is the sum of the series

$$(1-r)+(r-r^2)+(r^2-r^3)+\dots$$

when the series is convergent? [The series converges only if $-1<r\leqq 1$. Its sum is 1, except when $r=1$, when its sum is 0.]

6. Sum the series $r^2+\dfrac{r^2}{1+r^2}+\dfrac{r^2}{(1+r^2)^2}+\dots$. [The series is always convergent. Its sum is $1+r^2$, except when $r=0$, when its sum is 0.]

7. If we assume that $1+r+r^2+\dots$ is convergent we can prove that its sum is $1/(1-r)$ by means of § 70, (1) and (4). For if $1+r+r^2+\dots=s$,

$$s=1+r(1+r^2+\dots)=1+rs.$$

8. Sum the series $$r+\frac{r}{1+r}+\frac{r}{(1+r)^2}+\dots$$

when it is convergent. [The series is convergent if $-1<1/(1+r)<1$, i.e. if $r<-2$ or if $r>0$, and its sum is $1+r$. It is also convergent for $r=0$, when its sum is 0.]

9. Answer the same question for the series

$$r-\frac{r}{1+r}+\frac{r}{(1+r)^2}-\dots,\qquad r\pm\frac{r}{1-r}+\frac{r}{(1-r)^2}\pm\dots,$$

$$1\pm\frac{r}{1+r}+\left(\frac{r}{1+r}\right)^2\pm\dots,\qquad 1\pm\frac{r}{1-r}+\left(\frac{r}{1-r}\right)^2\pm\dots.$$

10. Consider the convergence of

$$(1+r)+(r^2+r^3)+\dots,\quad (1+r+r^2)+(r^3+r^4+r^5)+\dots,$$
$$1-2r+r^2+r^3-2r^4+r^5+\dots,\quad (1-2r+r^2)+(r^3-2r^4+r^5)+\dots,$$

and find their sums when they are convergent.

11. If a_n is positive and not greater than 1, the series $a_0+a_1r+a_2r^2+\dots$ is convergent for $0\leqq r<1$, and the sum of the series is not greater than $1/(1-r)$.

12. If in addition the series $a_0+a_1+a_2+\dots$ is convergent, the series $a_0+a_1r+a_2r^2+\dots$ is convergent for $0 \leqq r \leqq 1$, and its sum is not greater than the lesser of $a_0+a_1+a_2+\dots$ and $1/(1-r)$.

13. The series
$$1+\frac{1}{1.2}+\frac{1}{1.2.3}+\dots$$
is convergent. [For $1/(1.2\dots n)<1/2^{n-1}$.]

14. The series
$$1+\frac{1}{1.2}+\frac{1}{1.2.3.4}+\dots,\qquad 1+\frac{1}{1.2.3}+\frac{1}{1.2.3.4.5}+\dots$$
are convergent.

15. The general harmonic series
$$\frac{1}{a}+\frac{1}{a+b}+\frac{1}{a+2b}+\dots,$$
where a and b are positive, diverges to $+\infty$.

[For $u_n=1/(a+nb)>1/\{n(a+b)\}$. Now compare with $1+(1/2)+(1/3)+\dots$.]

16. Show that the series
$$(u_0-u_1)+(u_1-u_2)+(u_2-u_3)+\dots$$
is convergent if and only if u_n tends to a limit as $n \to \infty$.

17. If $u_1+u_2+u_3+\dots$ is divergent, so is any series formed by grouping the terms in brackets in any way to form new single terms.

18. Any series, formed by taking a selection of the terms of a convergent series of positive terms, is itself convergent.

72. The representation of functions of a continuous real variable by means of limits. In the preceding sections we have frequently been concerned with limits such as
$$\lim_{n\to\infty} \phi_n(x),$$
and series such as
$$u_1(x)+u_2(x)+\dots=\lim_{n\to\infty}\{u_1(x)+u_2(x)+\dots+u_n(x)\},$$
in which the function of n whose limit we are seeking involves, besides n, another variable x. In such cases the limit is of course a function of x. Thus in § 69 we came across the function
$$f(x)=\lim_{n\to\infty} n(\sqrt[n]{x}-1):$$
and the sum of the geometrical series $1+x+x^2+\dots$ is a function of x, viz. the function which is equal to $1/(1-x)$ if $-1<x<1$ and is undefined for all other values of x.

Many of the apparently 'arbitrary' or 'unnatural' functions considered in Ch. II are capable of a simple representation of this kind, as will appear from the following examples.

Examples XXXIII. 1. $\phi_n(x)=x$. Here n does not appear at all in the expression of $\phi_n(x)$, and $\phi(x)=\lim\phi_n(x)=x$ for all values of x.

2. $\phi_n(x)=x/n$. Here $\phi(x)=\lim\phi_n(x)=0$ for all values of x.

3. $\phi_n(x)=nx$. If $x>0$, $\phi_n(x)\to+\infty$; if $x<0$, $\phi_n(x)\to-\infty$: only for $x=0$ has $\phi_n(x)$ a limit (viz. 0) as $n\to\infty$. Thus $\phi(x)=0$ when $x=0$ and is not defined for any other value of x.

4. $\phi_n(x)=1/nx$, $nx/(nx+1)$.

5. $\phi_n(x)=x^n$. Here $\phi(x)=0$, $(-1<x<1)$; $\phi(x)=1$, $(x=1)$; and $\phi(x)$ is not defined for any other value of x.

6. $\phi_n(x)=x^n(1-x)$. Here $\phi(x)$ differs from the $\phi(x)$ of Ex. 5 in that it is defined and has the value 0 for $x=1$.

7. $\phi_n(x)=x^n/n$. Here $\phi(x)$ differs from the $\phi(x)$ of Ex. 6 in that it is defined and has the value 0 for $x=-1$ as well as $+1$.

8. $\phi_n(x)=x^n/(x^n+1)$. [$\phi(x)=0$, $(-1<x<1)$; $\phi(x)=\frac{1}{2}$, $(x=1)$; $\phi(x)=1$, $(x<-1$ or $x>1)$; and $\phi(x)$ is not defined for $x=-1$.]

9. $\phi_n(x)=x^n/(x^n-1)$, $1/(x^n+1)$, $1/(x^n-1)$, $1/(x^n+x^{-n})$, $1/(x^n-x^{-n})$.

10. $\phi_n(x)=(x^n-1)/(x^n+1)$, $(nx^n-1)/(nx^n+1)$, $(x^n-n)/(x^n+n)$. [In the first case $\phi(x)=1$ if $|x|>1$, $\phi(x)=-1$ if $|x|<1$, $\phi(x)=0$ if $x=1$ and $\phi(x)$ is not defined for $x=-1$. The second and third functions differ from the first in that they are defined both for $x=1$ and $x=-1$: the second has the value 1 and the third the value -1 for both these values of x.]

11. Construct an example in which $\phi(x)=1$, $(|x|>1)$; $\phi(x)=-1$, $(|x|<1)$; and $\phi(x)=0$, $(x=\pm1)$.

12. $\phi_n(x)=x\{(x^{2n}-1)/(x^{2n}+1)\}^2$, $n/(x^n+x^{-n}+n)$.

13. $\phi_n(x)=\{x^n f(x)+g(x)\}/(x^n+1)$. [Here $\phi(x)=f(x)$, $(|x|>1)$; $\phi(x)=g(x)$, $(|x|<1)$; $\phi(x)=\frac{1}{2}\{f(x)+g(x)\}$, $(x=1)$; and $\phi(x)$ is undefined for $x=-1$.]

14. $\phi_n(x)=(2/\pi)\arctan(nx)$. [$\phi(x)=1$, $(x>0)$; $\phi(x)=0$, $(x=0)$; $\phi(x)=-1$, $(x<0)$. This function is important in the Theory of Numbers, and is usually denoted by *sgn* x.]

15. $\phi_n(x)=(1/n)\sin nx\pi$. [$\phi(x)=0$ for all values of x.]

16. $\phi_n(x)=\sin nx\pi$. [$\phi(x)=0$ when x is an integer, and is otherwise undefined.]

17. $\phi_n(x)=(1/n)\cos nx\pi$, $\cos nx\pi$, $a\cos^2 nx\pi+b\sin^2 nx\pi$.

18. If $\phi_n(x)=\sin(n!x\pi)$, $\phi(x)=0$ for all *rational* values of x (Exs. XXVI. 9, XXVII. 8). The consideration of irrational values presents greater difficulties.

19. $\phi_n(x)=(\cos^2 x\pi)^n$. [$\phi(x)=0$ except when x is integral, when $\phi(x)=1$.]

20. $\phi_n(x)=(\sin^2 x\pi)^n$, $(\cos x\pi)^n$, $(\sin x\pi)^n$.

21. $\phi_n(x)=(a\cos^2 x\pi+b\sin^2 x\pi)^n$. [Here $\phi(x)=0$ if $|a\cos^2 x\pi+b\sin^2 x\pi|<1$, $\phi(x)=1$ if $a\cos^2 x\pi+b\sin^2 x\pi=1$, and $\phi(x)$ is otherwise undefined. For what values of x these respective conditions are satisfied depends on the values of a and b. Thus if a and b are both numerically less than unity, $\phi(x)=0$ for *all* values of x. Consider, e.g., the cases $a=b=1$; $a=b=\frac{1}{2}$; $a=b=2$; $a=1$, $b=2$; $a=2$, $b=1$; $a=2$, $b=\frac{1}{2}$.]

22. If $N \geqq 1752$, the number of days in the year N A.D. is

$$\lim\{365+(\cos^2\tfrac{1}{4}N\pi)^n-(\cos^2\tfrac{1}{100}N\pi)^n+(\cos^2\tfrac{1}{400}N\pi)^n\}.$$

73. Limits of Complex functions and series of Complex terms. In this chapter we have, up to the present, concerned ourselves only with real functions of n and series all of whose terms are real. There is however no difficulty in extending our ideas and definitions to the case in which the functions or the terms of the series are complex.

Suppose that $\phi(n)$ is complex and equal to

$$R(n)+iS(n),$$

where $R(n)$, $S(n)$ are real functions of n. Then *if, as $n \rightarrow \infty$, $R(n)$ and $S(n)$ converge respectively to limits r and s, we shall say that $\phi(n)$ converges to the limit $r+is$, and write*

$$\lim \phi(n)=r+is.$$

Similarly if u_n is complex and equal to v_n+iw_n we shall say that *the series*

$$u_1+u_2+u_3+\dots$$

is convergent and has the sum $r+is$, if the series

$$v_1+v_2+v_3+\dots, \quad w_1+w_2+w_3+\dots$$

are convergent and have the sums r, s respectively.

To say that $u_1+u_2+u_3+\dots$ is convergent and has the sum $r+is$ is of course the same as to say that the sum

$$s_n=u_1+u_2+\dots+u_n=(v_1+v_2+\dots+v_n)+i(w_1+w_2+\dots+w_n)$$

converges to the limit $r+is$ as $n\rightarrow\infty$.

In the case of real functions and series we also gave definitions of *divergence* and *oscillation* (finite or infinite). But in the case of complex functions and series there are so many possibilities—e.g. $R(n)$ may tend to $+\infty$ and $S(n)$ oscillate—that this is hardly worth while. When it is necessary to make further distinctions of this kind, we shall make them by stating the way in which the real or imaginary parts behave when taken separately.

74. The reader will find no difficulty in proving such theorems as the following, which are obvious extensions of theorems already proved for real functions and series.

(1) If $\lim \phi(n) = r + is$, then $\lim \phi(n+p) = r + is$, for any fixed value of p.

(2) If $u_1 + u_2 + \ldots$ is convergent and has the sum $r + is$, then $a + b + c + \ldots + k + u_1 + u_2 + \ldots$ is convergent and has the sum $a + b + c + \ldots + k + r + is$, and $u_{p+1} + u_{p+2} + \ldots$ is convergent and has the sum $r + is - u_1 - u_2 - \ldots - u_p$.

(3) If $\lim \phi(n) = a$ and $\lim \psi(n) = b$, then $\lim \{\phi(n) + \psi(n)\} = a + b$.

(4) If $\lim \phi(n) = a$, $\lim k\phi(n) = ka$.

(5) If $\lim \phi(n) = a$ and $\lim \psi(n) = b$, then $\lim \phi(n)\psi(n) = ab$.

(6) If $u_1 + u_2 + \ldots$ converges to the sum a, and $v_1 + v_2 + \ldots$ to the sum b, then $(u_1 + v_1) + (u_2 + v_2) + \ldots$ converges to the sum $a + b$.

(7) If $u_1 + u_2 + \ldots$ converges to the sum a, $ku_1 + ku_2 + \ldots$ converges to the sum ka.

(8) If $u_1 + u_2 + u_3 + \ldots$ is convergent, then $\lim u_n = 0$.

(9) If $u_1 + u_2 + u_3 + \ldots$ is convergent, so is any series formed by grouping the terms in brackets, and the sums of the two series are the same.

As an example, let us prove theorem (5). Let

$$\phi(n) = R(n) + iS(n), \quad \psi(n) = R'(n) + iS'(n), \quad a = r + is, \quad b = r' + is'.$$

Then $$R(n) \to r, \quad S(n) \to s, \quad R'(n) \to r', \quad S'(n) \to s'.$$

But $$\phi(n)\,\psi(n) = RR' - SS' + i(RS' + R'S)$$

and $$RR' - SS' \to rr' - ss', \quad RS' + R'S \to rs' + r's,$$

so that $$\phi(n)\,\psi(n) \to rr' - ss' + i(rs' + r's),$$

i.e. $$\phi(n)\,\psi(n) \to (r + is)(r' + is') = ab.$$

The following theorems are of a somewhat different character.

(10) In order that $\phi(n) = R(n) + iS(n)$ should converge to zero as $n \to \infty$ it is necessary and sufficient that

$$|\phi(n)| = \sqrt{\{R(n)\}^2 + \{S(n)\}^2}$$

should converge to zero.

If $R(n)$ and $S(n)$ both converge to zero it is plain that $\sqrt{(R^2 + S^2)}$ does so. The converse follows from the fact that the numerical value of R or S cannot be greater than $\sqrt{(R^2 + S^2)}$.

(11) More generally, in order that $\phi(n)$ should converge to a limit l it is necessary and sufficient that

$$|\phi(n)-l|$$

should converge to zero.

For $\phi(n)-l$ converges to zero, and we can apply (10).

75. The limit of x^n as $n \to \infty$, x being any complex number. Let us consider the important case in which $\phi(n)=x^n$. This problem has already been discussed for real values of x in § 66.

If $x^n \to l$, $x^{n+1} \to l$, by (1) above. But since $x^{n+1}=x \,.\, x^n$, $x^{n+1} \to xl$, by (4) above; and therefore $l=xl$, which is only possible if (*a*) $l=0$ or (*b*) $x=1$. If $x=1$, $\lim x^n=1$. Apart from this special case the limit, if it exists, can only be zero.

Now if
$$x=r(\cos\theta+i\sin\theta),$$
where r is positive, we know that
$$x^n=r^n(\cos n\theta+i\sin n\theta),$$
so that $|x^n|=r^n$. Thus $|x^n|$ tends to zero if and only if $r<1$; and it follows from (10) of the last paragraph that
$$\lim x^n=0,$$
if and only if $r<1$. In no other case does x^n converge to a limit, except when $x=1$ and $x^n \to 1$.

76. The geometric series $1+x+x^2+\dots$, when x is complex. Since
$$s_n=1+x+x^2+\dots+x^{n-1}=(1-x^n)/(1-x),$$
unless $n=1$, when the value of s_n is n, it follows that *the series* $1+x+x^2+\dots$ *is convergent if and only if* $r=|x|<1$. *And its sum when convergent is* $1/(1-x)$.

Thus if $x=r(\cos\theta+i\sin\theta)=r\operatorname{Cis}\theta$, and $r<1$,
$$1+x+x^2+\dots=1/(1-r\operatorname{Cis}\theta),$$
or
$$\begin{aligned}1+r\operatorname{Cis}\theta+r^2\operatorname{Cis}2\theta+\dots &= 1/(1-r\operatorname{Cis}\theta)\\ &=(1-r\cos\theta+ir\sin\theta)/(1-2r\cos\theta+r^2),\end{aligned}$$
or, separating the real and imaginary parts,
$$1+r\cos\theta+r^2\cos 2\theta+\dots=(1-r\cos\theta)/(1-2r\cos\theta+r^2),$$
$$r\sin\theta+r^2\sin 2\theta+\dots=r\sin\theta/(1-2r\cos\theta+r^2),$$
provided $r<1$. If we change θ into $\theta+\pi$ we see that these results hold also for negative values of r numerically less than 1. Thus they hold for $-1<r<1$.

Examples XXXIV. 1. Prove directly that $\phi(n) = r^n \cos n\theta$ converges to 0 if $r<1$ and to 1 if $r=1$, $\theta=0$. Prove further that if $r=1$, $\theta \neq 0$ it oscillates finitely, if $r>1$, $\theta=0$ it diverges to $+\infty$ and if $r>1$, $\theta \neq 0$ it oscillates infinitely.

2. Establish a similar series of results for $\phi(n) = r^n \sin n\theta$.

3. Prove (as for the case of a real x in Ex. XXXII. 7) that if $1+x+x^2+\ldots$ converges its sum can only be $1/(1-x)$.

4. Prove that
$$x^m + x^{m+1} + \ldots = x^m/(1-x),$$
$$x^m - x^{m+1} + \ldots = x^m/(1+x),$$
$$x^m + 2x^{m+1} + 2x^{m+2} + \ldots = x^m(1+x)/(1-x),$$
$$x^m - 2x^{m+1} + 2x^{m+2} - \ldots = x^m(1-x)/(1+x),$$
if and only if $|x|<1$. Which of the theorems of § 74 do you use?

5. Let (in the notation of Chap. III, §§ 25 *et seq.*) $\overline{P_0P_1}=1$, $\overline{P_1P_2}=x$, $\overline{P_2P_3}=x^2\ldots$ where $x = r\,\mathrm{Cis}\,\theta$. Plot the points P_0, P_1, P_2, ..., and show how the figure obtained indicates the result of § 76. Prove that, if $r<1$, the point P_n, where n is large, is very near to the point
$$(1-r\cos\theta)/(1-2r\cos\theta+r^2), \quad r\sin\theta/(1-2r\cos\theta+r^2).$$

6. Prove that, if $-1<r<1$,
$$1+2r\cos\theta+2r^2\cos 2\theta+\ldots=(1-r^2)/(1-2r\cos\theta+r^2).$$

7. The series $\quad 1+\{x/(1+x)\}+\{x/(1+x)\}^2+\ldots$
converges to the sum $1\Big/\left(1-\dfrac{x}{1+x}\right)=1+x$ if $|x/(1+x)|<1$. Show that this is equivalent to the assertion that x has a real part greater than $-\frac{1}{2}$.

8. Determine similarly the regions of values of x for which the series, obtained by writing x for r in Ex. XXXII. 9, are convergent, and find their sums when they are convergent.

MISCELLANEOUS EXAMPLES ON CHAPTER IV.

1. The function $\phi(n)$ takes for $n=0, 1, 2, \ldots$ the values 1, 0, 0, 0, 1, 0, 0, 0, 1, ... Express $\phi(n)$ in terms of n by a formula which does not involve trigonometrical functions. [$\phi(n)=\frac{1}{4}\{1+(-1)^n+i^n+(-i)^n\}$.]

2. If $\phi(n)$ steadily increases, and $\psi(n)$ steadily decreases, as n tends to ∞, and if $\psi(n)>\phi(n)$ for all values of n, then both $\phi(n)$ and $\psi(n)$ tend to limits, and $\lim\phi(n) \leqq \lim\psi(n)$. [This is an intermediate corollary from § 64.]

3. Prove that if
$$\phi(n)=\left(1+\frac{1}{n}\right)^n, \quad \psi(n)=\left(1-\frac{1}{n}\right)^{-n},$$
then $\phi(n+1)>\phi(n)$ and $\psi(n+1)<\psi(n)$.

[The first inequality has already been proved in § 67; the second may be proved similarly.]

4. Prove also that $\psi(n)>\phi(n)$ for all values of n: and deduce (by means of the preceding examples) that both $\phi(n)$ and $\psi(n)$ tend to limits as n tends to ∞ *.

5. If $\binom{m}{n}=\dfrac{m(m-1)\dots(m-n+1)}{1.2\dots n}$, m not being a positive integer, and $-1<x<1$, then $u_n=\binom{m}{n}x^n\rightarrow 0$ as $n\rightarrow\infty$.

[For $u_{n+1}/u_n=\{(m-n)/(n+1)\}x\rightarrow -x$. Now apply Ex. XXX. 6.]

6. The arithmetic mean of the products of all distinct pairs of positive integers, whose sum is n, is denoted by S_n. Show that $\lim(S_n/n^2)=1/6$. (*Math. Trip.* 1903.)

7. If $x_1=\frac{1}{2}\{x+(A/x)\}$, $x_2=\frac{1}{2}\{x_1+(A/x_1)\}$, and so on, x and A being positive, prove that $\lim x_n=\sqrt{A}$.

$\left[\text{Prove first that } \dfrac{x_n-\sqrt{A}}{x_n+\sqrt{A}}=\left(\dfrac{x-\sqrt{A}}{x+\sqrt{A}}\right)^{2^n}.\right]$

8. If $\phi(n)$ is a positive integer for all values of n, and tends to ∞ with n, then $x^{\phi(n)}\rightarrow 0$ or $+\infty$ according as $0<x<1$ or $x>1$. Discuss the behaviour of $x^{\phi(n)}$, as $n\rightarrow\infty$, for other values of x.

9†. If a_n increases (decreases) steadily as n increases, the same is true of $(a_1+a_2+\dots+a_n)/n$.

10. If $x_{n+1}=\sqrt{(k+x_n)}$, and k and x_1 are positive, the sequence x_1, x_2, x_3, ... is an increasing or decreasing sequence according as x_1 is less than or greater than a, the positive root of the equation $x^2=x+k$; and in either case $x_n\rightarrow a$ as $n\rightarrow\infty$.

11. If $x_{n+1}=k/(1+x_n)$, and k and x_1 are positive, the sequence x_1, x_2, x_3, ... is an increasing or decreasing sequence, according as x_1 is less than or greater than a, the positive root of the equation $x^2+x=k$; and in either case $x_n\rightarrow a$ as $n\rightarrow\infty$.

12. Suppose that $f(x)$ is a positive and increasing function of x such that the equation $x=f(x)$ has just one positive root a. Show, graphically or otherwise, that if $x_1>0$ and $x_{n+1}=f(x_n)$ then the sequence x_1, x_2, ... has the limit a as $n\rightarrow\infty$.

Discuss the case in which the equation $x=f(x)$ has several positive roots.

13. If x_1, x_2 are positive and $x_{n+1}=\frac{1}{2}(x_n+x_{n-1})$, the sequences x_1, x_3, x_5, ... and x_2, x_4, x_6, ... are one a decreasing and one an increasing sequence, and their common limit is $\frac{1}{3}(a_1+2a_2)$.

14. Draw a graph of the function y defined by the equation

$$y=\lim_{n\rightarrow\infty}\frac{x^{2n}\sin\frac{1}{2}\pi x+x^2}{x^{2n}+1}.$$

(*Math. Trip.* 1901.)

* A proof that $\lim\{\psi(n)-\phi(n)\}=0$, and that therefore each function tends to the limit e, will be found in Chrystal's *Algebra*, vol. ii, p. 78. We shall however prove this in Ch. IX by a different method.

† Exs. 9—13 are taken from Bromwich's *Infinite Series*.

15. The function $$y=\lim_{n\to\infty}\frac{1}{1+n\sin^2\pi x}$$
is equal to 0 except when x is an integer, and then equal to 1. The function
$$y=\lim_{n\to\infty}\frac{\psi(x)+n\phi(x)\sin^2\pi x}{1+n\sin^2\pi x}$$
is equal to $\phi(x)$ unless x is an integer, and then equal to $\psi(x)$.

16. Show that the graph of the function
$$y=\lim_{n\to\infty}\frac{x^n\phi(x)+x^{-n}\psi(x)}{x^n+x^{-n}}$$
is composed of parts of the graphs of $\phi(x)$ and $\psi(x)$, together with (as a rule) one isolated point. Is y defined for (*a*) $x=1$, (*b*) $x=-1$, (*c*) $x=0$?

17. Prove that the function y, which is equal to 0 when x is rational and to 1 when x is irrational, may be represented in the form
$$y=\lim_{m\to\infty} sgn\{\sin^2(m!\pi x)\}$$
where, as in Ex. XXXIII. 14, $sgn\, z=\lim_{n\to\infty}(2/\pi)\arctan(nz)$.

[If x is rational, $\sin^2(m!\pi x)$, and therefore $sgn\{\sin^2(m!\pi x)\}$ is equal to zero from a certain value of m onwards: if x is irrational, $\sin^2(m!\pi x)$ is always positive, and so $sgn\{\sin^2(m!\pi x)\}$ is always equal to 1.]

Prove that y may also be represented in the form
$$1-\lim_{m\to\infty}[\lim_{n\to\infty}\{\cos(m!\pi x)\}^{2n}].$$

18. Sum the series
$$\sum_1^\infty\frac{1}{\nu(\nu+1)},\quad \sum_1^\infty\frac{1}{\nu(\nu+1)\dots(\nu+k)}.$$
[Since
$$\frac{1}{\nu(\nu+1)\dots(\nu+k)}=\frac{1}{k}\left\{\frac{1}{\nu(\nu+1)\dots(\nu+k-1)}-\frac{1}{(\nu+1)(\nu+2)\dots(\nu+k)}\right\},$$
we find $$\sum_1^n\frac{1}{\nu(\nu+1)\dots(\nu+k)}=\frac{1}{k}\left\{\frac{1}{1.2\dots k}-\frac{1}{(n+1)(n+2)\dots(n+k)}\right\}$$
and so $$\sum_1^\infty\frac{1}{\nu(\nu+1)\dots(\nu+k)}=\frac{1}{k(k!)}.]$$

19. If $|x|<|a|$, $$\frac{L}{x-a}=-\frac{L}{a}\left(1+\frac{x}{a}+\frac{x^2}{a^2}+\dots\right);$$
while if $|x|>|a|$, $$\frac{L}{x-a}=\frac{L}{x}\left(1+\frac{a}{x}+\frac{a^2}{x^2}+\dots\right).$$

20. **Expansion of $(Ax+B)/(ax^2+2bx+c)$ in powers of x.** Let α, β be the roots of $ax^2+2bx+c=0$, so that $ax^2+2bx+c\equiv a(x-\alpha)(x-\beta)$. It is easy to verify that (unless $\alpha=\beta$)
$$\frac{Ax+B}{ax^2+2bx+c}\equiv\frac{1}{a(\alpha-\beta)}\left(\frac{A\alpha+B}{x-\alpha}-\frac{A\beta+B}{x-\beta}\right).$$

We shall suppose A, B, a, b, c all real. Then there are two cases, according as $b^2 \gtrless ac$.

(1) If $b^2 > ac$, the roots α, β are real and distinct. If $|x|$ is less than either $|\alpha|$ or $|\beta|$ we can expand $1/(x-\alpha)$ and $1/(x-\beta)$ in ascending powers of x (Ex. 19). If $|x|$ is greater than either $|\alpha|$ or $|\beta|$ we must expand in descending powers of x; while if $|x|$ lies between $|\alpha|$ and $|\beta|$ one fraction must be expanded in ascending and one in descending powers of x. The reader should write down the actual results. If $|x|$ is equal to $|\alpha|$ or $|\beta|$ no such expansion is possible.

(2) If $b^2 < ac$ the roots are conjugate complex numbers (Chap. III, § 34) and we can write

$$\alpha = \rho \operatorname{Cis} \phi, \qquad \beta = \rho \operatorname{Cis}(-\phi),$$

where $\rho^2 = \alpha\beta = c/a$, $\rho\cos\phi = \frac{1}{2}(\alpha+\beta) = -b/a$, so that $\cos\phi = -\sqrt{(b^2/ac)}$, $\sin\phi = \sqrt{\{1-(b^2/ac)\}}$.

If $|x| < \rho$ each fraction may be expanded in ascending powers of x. The coefficient of x^n will be found to be

$$\{A\rho \sin n\phi + B\sin(n+1)\phi\}/a\rho^{n+1}\sin\phi.$$

If $|x| > \rho$ we obtain a similar expansion in descending powers, while if $|x| = \rho$ no such expansion is possible.

21. Show that, if $|x| < 1$,

$$1 + 2x + 3x^2 + \dots + (n+1)x^n + \dots = 1/(1-x)^2.$$

$\left[\text{The sum to } n \text{ terms is } \dfrac{1-x^n}{(1-x)^2} - \dfrac{nx^n}{1-x}.\right]$

22. Expand $L/(x-a)^2$ in powers of x, ascending or descending according as $|x| < |a|$ or $|x| > |a|$.

23. Show that if $b^2 = ac$ and $|ax| < |b|$

$$(Ax+B)/(ax^2+2bx+c) = \sum_1^\infty p_n x^n,$$

where $p_n = \{(-a)^n/b^{n+2}\}\{(n+1)aB - nbA\}$, and find the corresponding expansion, in descending powers of x, which holds when $|ax| > |b|$.

24. Verify the result of Ex. 20 in the case of the fraction $1/(1+x^2)$. [We have $1/(1+x^2) = \Sigma x^n \sin\{\frac{1}{2}(n+1)\pi\} = 1 - x^2 + x^4 - \dots$.]

25. Prove that, if $|x| < 1$, $1/(1+x+x^2) = (2/\sqrt{3})\sum_0^\infty x^n \sin\{\frac{2}{3}(n+1)\pi\}$.

26. Expand $(1+x)/(1+x^2)$, $(1+x^2)/(1+x^3)$ and $(1+x+x^2)/(1+x^4)$ in ascending powers of x. For what values of x do your results hold?

27. If $a/(a+bx+cx^2) = 1 + p_1x + p_2x^2 + \dots$ then

$$1 + p_1{}^2x + p_2{}^2x^2 + \dots = \frac{a+cx}{a-cx}\,\frac{a^2}{a^2-(b^2-2ac)x+c^2x^2}.$$

(*Math. Trip.* 1900.)

28. If $\lim\limits_{x\to\infty} s_n = l$, then

$$\lim_{n\to\infty} (s_1+s_2+\ldots+s_n)/n = l.$$

[Let $s_n = l + t_n$. Then we have to prove that $(t_1+t_2+\ldots+t_n)/n$ tends to zero if t_n does so.

We divide the numbers $t_1, t_2, \ldots t_n$ into two sets $t_1, t_2, \ldots, t_p$ and $t_{p+1}, t_{p+2}, \ldots, t_n$. Here we suppose that p is a function of n which tends to ∞ with n, but *more slowly than* n, so that $p\to\infty$ but $p/n\to 0$: e.g. we might suppose p to be the integral part of $\sqrt{n}$.

Let ϵ be any positive number. However small ϵ may be, we can choose n_0 so that $t_{p+1}, t_{p+2}, \ldots, t_n$ are all numerically less than $\frac{1}{2}\epsilon$ when $n \geqq n_0$, and so

$$|(t_{p+1}+t_{p+2}+\ldots+t_n)/n| < \tfrac{1}{2}\epsilon\,(n-p)/n < \tfrac{1}{2}\epsilon.$$

But, if A is the greatest of the moduli of all the numbers $t_1, t_2, \ldots$, we have also

$$|(t_1+t_2+\ldots+t_p)/n| < pA/n,$$

and, if n_0 is large enough, this will also be less than $\frac{1}{2}\epsilon$ when $n \geqq n_0$, since $p/n\to 0$ as $n\to\infty$. Thus, if n_0 is large enough,

$$|(t_1+t_2+\ldots+t_n)/n| \leqq |(t_1+t_2+\ldots+t_p)/n| + |(t_{p+1}+\ldots+t_n)/n| < \epsilon,$$

when $n \geqq n_0$: which proves the theorem.

The reader, if he desires to become expert in dealing with questions about limits, should study the argument here given with great care. It is very often necessary, in proving the limit of some given expression to be zero, to split it into two parts which have to be proved to have the limit zero in slightly different ways. When this is the case the proof is never very easy.

The point of the proof is this: we have to prove that $(t_1+t_2+\ldots+t_n)/n$ is small when n is large, the t's being small when their suffixes are large. We split up the terms in the bracket into two groups. The terms in the first group are not all small, but their *number* is small compared with n. The number in the second group is *not* small compared with n, but the terms are all small, and their number at any rate less than n, so that their sum is small compared with n. Hence each of the parts into which $(t_1+t_2+\ldots+t_n)/n$ has been divided is small when n is large.]

29. If $\phi(n)-\phi(n-1)\to l$ as $n\to\infty$ then also $\phi(n)/n\to l$.

[If we put $\phi(n) = s_1+s_2+\ldots+s_n$, we have $\phi(n)-\phi(n-1) = s_n$, and the theorem reduces to that proved in the last example.]

30. If $s_n = \frac{1}{2}\{1-(-1)^n\}$, so that s_n is equal to 1 or 0 according as n is odd or even, then $(s_1+s_2+\ldots+s_n)/n\to\frac{1}{2}$ as $n\to\infty$.

[This example proves that the converse of 28 is *not* true: for s_n oscillates as $n\to\infty$.]

31. Let c_n, s_n denote the sums of the first n terms of the series

$$\tfrac{1}{2}+\cos\theta+\cos 2\theta+\ldots, \quad \sin\theta+\sin 2\theta+\ldots.$$

Prove that

$$\lim\,(c_1+c_2+\ldots+c_n)/n = 0, \quad \lim\,(s_1+s_2+\ldots+s_n)/n = \tfrac{1}{2}\cot\tfrac{1}{2}\theta.$$

CHAPTER V.

LIMITS OF FUNCTIONS OF A CONTINUOUS VARIABLE. CONTINUOUS AND DISCONTINUOUS FUNCTIONS.

77. Limits as x tends to ∞. We shall now return to functions of a continuous real variable. We shall denote the typical such function by $\phi(x)$. We suppose x to assume successively all values corresponding to points on our fundamental straight line L, starting from some definite point on the line and progressing always to the right. This variation of x, like the corresponding variation of n (Chap. IV, § 48), is often conveniently thought of as taking place in time. In these circumstances we say that *x tends to ∞*, and write $x \rightarrow \infty$. The only difference between the 'tending of n to ∞' discussed in the last chapter, and this 'tending of x to ∞,' is that x varies through all values as it tends to ∞, i.e. that the point P which corresponds to x coincides in turn with every point of L to the right of its initial position, whereas n tends to ∞ by a series of jumps. We can express this distinction by saying that x tends *continuously* to ∞.

As we explained at the beginning of the last chapter, there is a very close correspondence between functions of x and functions of n. Every function of n may be regarded as a selection from the values of a function of x. In the last chapter we discussed the peculiarities which may characterise the behaviour of a function $\phi(n)$ as n tends to ∞. Now we are concerned with the same problem for a function $\phi(x)$: and the definitions and theorems to which we are led are practically repetitions of those of the last chapter. Thus corresponding to Def. 1 of § 51 we have:

DEFINITION 1. *The function $\phi(x)$ is said to tend to the limit l as x tends to ∞ if, when any positive number ϵ, however small, is*

assigned, a value X can be chosen such that, for all values of x equal to or greater than X, $\phi(x)$ differs from l by less than ϵ, i.e. if

$$|\phi(x) - l| < \epsilon \qquad (x \geqq X).$$

When this is the case we may write

$$\lim_{x \to \infty} \phi(x) = l, \quad \underset{(x \to \infty)}{\phi(x) \to l},$$

or, when there is no risk of ambiguity, simply $\lim \phi(x) = l$, or $\phi(x) \to l$. Similarly we have:

DEFINITION 2. *The function $\phi(x)$ is said to tend to $+\infty$ with x if when any number G, however large, is assigned, we can choose X so that*

$$\phi(x) > G \qquad (x \geqq X).$$

We then write $\qquad \phi(x) \to +\infty.$

Similarly we define $\phi(x) \to -\infty$. Finally we have:

DEFINITION 3. *If the conditions of none of the two preceding definitions are satisfied $\phi(x)$ is said to oscillate as x tends to ∞. If, for all values of x, $|\phi(x)|$ is less than some constant K, $\phi(x)$ is said to oscillate finitely; otherwise infinitely.*

The reader will remember that in the last chapter we considered very carefully various less formal ways of expressing the facts represented by the equations $\phi(n) \to l$, $\phi(n) \to +\infty$. Similar modes of expression may of course be used in the present case. Thus we may say that $\phi(x)$ is small or nearly equal to l or large when n is large, using the words 'small,' 'nearly,' 'large' in a sense precisely similar to that in which they were used in Ch. IV.

Examples XXXV. 1. Consider the behaviour of the following functions as $x \to \infty$: $(1/x)$, $1 + (1/x)$, x^2, x^k, $[x]$, $x - [x]$, $[x] + \sqrt{\{x - [x]\}}$.

The first four functions correspond exactly to functions of n fully discussed in Ch. IV. The graphs of the last three were constructed in Ch. II. (Exs. XVII.), and the reader will see at once that $[x] \to +\infty$, $x - [x]$ oscillates finitely, and $[x] + \sqrt{\{x - [x]\}} \to +\infty$.

One simple remark may be inserted here. The function $\phi(x) = x - [x]$ oscillates (between 0 and 1) as is obvious from the form of its graph. It is equal to zero whenever x is an integer, so that the function $\phi(n)$ derived from it is always zero and so tends to the limit zero. The same is true of

$$\phi(x) = \sin x\pi, \quad \phi(n) = \sin n\pi = 0.$$

In such cases as these, it is evident that $\phi(x)\to l$ or $\phi(x)\to +\infty$ or $-\infty$ involves the corresponding property for $\phi(n)$, but that the converse is by no means true.

2. Consider in the same way the functions:

$\cos x\pi$, $\tan x\pi$, $(\cos x\pi)/x$, $(\tan x\pi)/x$, $(1/x)+\cos x\pi$, $x\cos x\pi$, $x^2\cos x\pi$, $x\cos^2 x\pi$, $(x\cos x\pi)^2$, $a\cos^2 x\pi+b\sin^2 x\pi$, $(a\cos^2 x\pi+b\sin^2 x\pi)/x$,

illustrating your remarks by means of the graphs of the functions*.

3. Give a geometrical explanation of Def. 1, analogous to the geometrical explanation of Ch. IV, § 52.

4. If $\phi(x)\to l$, $\phi(x)\cos x\pi$ and $\phi(x)\sin x\pi$ oscillate finitely. If $\phi(x)\to +\infty$ (or $-\infty$) they oscillate infinitely. The graph of either function is a wavy curve oscillating between the curves $y=\phi(x)$, $y=-\phi(x)$.

5. Discuss the behaviour, as $x\to\infty$, of the function

$$y=f(x)\cos^2 x\pi+F(x)\sin^2 x\pi.$$

The graph of y is a curve oscillating between the curves $y=f(x)$, $y=F(x)$. Consider in particular the cases

(i) $f(x)=1+(1/x)$, $F(x)=1-(1/x)$; (ii) $f(x)=a+(a/x)$, $F(x)=b+(\beta/x)$, where $a\neq b$; (iii) $f(x)=1$, $F(x)=x$; (iv) $f(x)=-x$, $F(x)=x$; (v) $f(x)=\sin x\pi$, $F(x)=\cos x\pi$; (vi) $f(x)=\cos^4 x\pi+3\sin^4 x\pi$, $F(x)=3\cos^4 x\pi+\sin^4 x\pi$.

78. Limits as x tends to $-\infty$. The reader will have no difficulty in finding for himself definitions of the meaning of the assertions 'x tends to $-\infty$' ($x\to-\infty$) and

$$\lim_{x\to-\infty}\phi(x)=l,\qquad \underset{(x\to-\infty)}{\phi(x)\to\infty}\ (\text{or } -\infty).$$

In fact if $x=-y$ and $\phi(x)=\phi(-y)=\psi(y)$, then x tends to $-\infty$ as y tends to ∞, and the question of the behaviour of $\phi(x)$ as x tends to $-\infty$ is the same as that of the behaviour of $\psi(y)$ as y tends to ∞.

79. Theorems corresponding to those of Ch. IV, §§ 56—63. The theorems concerning the sums, products, and quotients of functions, proved in Ch. IV, are all true (with the obvious verbal alterations which the reader will have no difficulty in supplying) for functions of the continuous variable x. Not only the enunciations but the proofs remain substantially the same.

Ex. Draw up a table, with examples of each case, similar to the table on p. 130 in Ch. IV, i.e. to illustrate the different possibilities with regard to the behaviour of $\phi(x)+\psi(x)$ when the behaviour of $\phi(x)$ and $\psi(x)$ is known. The other tables of Ch. IV suggest similar examples.

* The reader has probably already drawn graphs of some of the functions considered in Exs. 2, 4, 5, while engaged on Ch. II and in particular Exs. xvi.

80. Steadily increasing or decreasing functions. The definition which corresponds to that of § 64 is as follows: *the function $\phi(x)$ will be said to increase steadily with x if $\phi(x_1) \geqq \phi(x_0)$ whenever $x_1 > x_0$.* In many cases, of course, this condition is only satisfied from a definite value $x = X$ onwards, i.e. when $x_1 > x_0 \geqq X$.

The theorem which follows requires no alteration but that of n into x: and the proof is also practically the same.

The reader should consider whether or no the following functions increase steadily with x (or at any rate increase steadily from a certain value of x onwards): $x^2 - x$, $x + \sin x$, $x + 2\sin x$, $x^2 + 2\sin x$, $[x]$, $[x] + \sin x$, $[x] + \sqrt{\{x - [x]\}}$. All these functions tend to $+\infty$ with x.

Ex. Show that if $\phi(x)$ steadily increases (or decreases) as $x \to \infty$, then the behaviour of $\phi(x)$ as $x \to \infty$ is the same as that of $\phi(n)$ as $n \to \infty$.

81. Limits as x tends to 0. Let $\phi(x)$ be such a function of x that $\lim\limits_{x\to\infty} \phi(x) = l$, and let $y = 1/x$. Then

$$\phi(x) = \phi(1/y) = \psi(y)$$

say. As x tends to ∞, y tends to the limit 0, and $\psi(y)$ tends to the limit l.

Let us now dismiss x and consider $\psi(y)$ simply as a function of y. We are for the moment concerned only with those values of y which correspond to *large* positive values of x, that is to say with *small* positive values of y. And $\psi(y)$ has the property that by making y sufficiently small we can make $\psi(y)$ differ by as little as we please from l. To put the matter more precisely, the statement expressed by $\lim \phi(x) = l$ meant that, when any positive number ϵ, however small, was assigned, we could choose X so that $|\phi(x) - l| < \epsilon$ for all values of x greater than or equal to X. But this is the same thing as saying that we can choose $\eta = 1/X$ so that $|\psi(y) - l| < \epsilon$ for all positive values of y less than or equal to η.

We are thus led to the following definitions.

A. *If when any positive number ϵ, however small, is assigned we can choose η so that*

$$|\phi(y) - l| < \epsilon,$$

for $0 < y \leqq \eta$, we say that $\phi(y)$ tends to the limit l when y tends to 0 by positive values, and we write

$$\lim_{y\to+0} \phi(y) = l.$$

B. *If when any number G, however large, is assigned we can choose η so that*

$$\phi(y) > G$$

for $0 < y \leqq \eta$, we say that $\phi(y)$ tends to $+\infty$ as y tends to 0 by positive values, and we write

$$\underset{(y\to+0)}{\phi(y)} \to +\infty.$$

We define in a similar way the meaning of '$\phi(y)$ tends to the limit l as y tends to 0 by negative values, or $\lim\limits_{y\to-0} \phi(y) = l$.' We have in fact only to alter $0 < y \leqq \eta$ to $-\eta \leqq y < 0$ in A. The reader will find it a useful exercise to write out formal definitions of the statements expressed by

$$\text{`}\underset{(y\to-0)}{\phi(y)} \to +\infty,\text{'} \qquad \text{`}\underset{(y\to+0)}{\phi(y)} \to -\infty,\text{'} \qquad \text{`}\underset{(y\to-0)}{\phi(y)} \to -\infty.\text{'}$$

If $\lim\limits_{(y\to+0)} \phi(y) = l$ **and** $\lim\limits_{(y\to-0)} \phi(y) = l$, we write simply

$$\lim_{(y\to0)} \phi(y) = l.$$

This case is so important that it is worth while to give a formal definition.

If when any positive number ϵ, however small, is assigned we can choose η so that, for all values of y different from zero but **numerically** *less than or equal to η, $\phi(y)$ differs from l by less than ϵ, we say that $\phi(y)$ tends to the limit l as y tends to 0, and write*

$$\lim_{y\to0} \phi(y) = l.$$

So also if $\underset{(y\to+0)}{\phi(y)} \to +\infty$ and $\underset{(y\to-0)}{\phi(y)} \to +\infty$ we write $\underset{y\to0}{\phi(y)} \to +\infty$. Similarly we define the statement $\underset{(y\to0)}{\phi(y)} \to -\infty$. Finally, if $\phi(y)$ does not tend to a limit, or to $+\infty$, or to $-\infty$, as $y \to 0$, we say that $\phi(y)$ oscillates as $y \to 0$, finitely or infinitely as the case may be.

The preceding definitions have been stated in terms of a variable denoted by y: what letter is used is of course immaterial, and we may suppose x written instead of y throughout them.

82. Limits as x tends to a. Suppose that $\lim\limits_{y\to0} \phi(y) = l$ and write

$$y = x - a, \quad \phi(y) = \phi(x-a) = \psi(x).$$

As $y \to 0$, $x \to a$ and $\psi(x) \to l$, and we are naturally led to write

$$\lim_{x \to a} \psi(x) = l, \qquad \psi(x) \underset{x \to a}{\to} l$$

or simply $\lim \psi(x) = l$ or $\psi(x) \to l$, and to say that $\psi(x)$ *tends to the limit* l *as* x *tends to* a. The meaning of this equation may be directly and formally defined as follows: *if, given* ϵ, *we can always determine* η *so that*

$$|\phi(x) - l| < \epsilon$$

for all values of x *such that* $0 < |x - a| \leqq \eta$, *then*

$$\lim_{x \to a} \phi(x) = l.$$

In other words, given any positive number ϵ we can find another η such that if x is different from a, but its difference from a is less than η, $\phi(x)$ will differ from l by less than ϵ.

By restricting ourselves to values of x greater than a, i.e. by replacing $0 < |x - a| \leqq \eta$ by $a < x \leqq a + \eta$, we define '$\phi(x)$ tends to l when x approaches a from the right'; which we may write as

$$\lim_{(x \to a+0)} \phi(x) = l.$$

Similarly we define

$$\lim_{x \to a-0} \phi(x) = l.$$

Thus $\lim_{x \to a} \phi(x) = l$ is equivalent to the two assertions

$$\lim_{x \to a+0} \phi(x) = l = \lim_{x \to a-0} \phi(x).$$

And we can give similar definitions referring to the cases in which $\phi(x) \to +\infty$ (or $-\infty$) as $x \to a$ through values greater (or less) than a; but it is probably unnecessary to dwell further on these definitions, since they are exactly similar to those stated above in the special case when $a = 0$, and since we can always discuss the behaviour of $\phi(x)$ as $x \to a$ by putting $x = y + a$ and supposing that $y \to 0$.

Examples XXXVI. 1. If

$$\phi(x) \underset{x \to a}{\to} b, \quad \psi(x) \underset{x \to a}{\to} c,$$

then $\phi(x) \pm \psi(x) \to b \pm c$, $\phi(x)\psi(x) \to bc$, and $\phi(x)/\psi(x) \to b/c$, unless in the last case $c = 0$.

[We saw in § 79 that the theorems of Ch. IV, §§ 56 *et seq.* held also for functions of x when $x \to \infty$ (or $-\infty$). By putting $x = 1/y$ we may extend

them to functions of y, when $y \to 0$, and by putting $y=x+a$ to functions of x, when $x \to a$.

The reader should however try to prove them directly from the formal definition given above. Thus, in order to obtain a strict direct proof of the first result he need only take the proof of Theorem I. in Ch. IV. and write throughout x for n, a for ∞ and $0<|x-a| \leqq \eta$ instead of $n \geqq n_0$.]

2. If m is a positive integer $x^m \to 0$ as $x \to 0$.

3. If m is a negative integer $x^m \to +\infty$ as $x \to +0$, while $x^m \to -\infty$ or $+\infty$ as $x \to -0$, according as m is odd or even. If $m=0$, $x^m=1$ and $x^m \to 1$.

4. $\lim\limits_{x \to 0} (a+bx+cx^2+ \dots +kx^m)=a$.

5. $\lim\limits_{x \to 0} \{(a+bx+ \dots +kx^m)/(\alpha+\beta x+ \dots +\kappa x^\mu)\}=a/\alpha$ unless $\alpha=0$. If $\alpha=0$ the function tends to $+\infty$ or $-\infty$, as $x \to +0$, according as a, β have like or unlike signs; the case is reversed if $x \to -0$.

6. $\lim\limits_{x \to a} x^m=a^m$, if m is any positive or negative integer.

[If $m>0$, put $x=y+a$ and apply Ex. 4. When $m<0$ the result follows from the theorem concerning $1/\phi(x)$. There is one exceptional case, viz. when $a=0$ and m is negative.

It follows at once that if $P(x)$ is any polynomial, $\lim P(x)=P(a)$.]

7. $\lim\limits_{x \to a} R(x)=R(a)$, if R denotes any rational function and a is not one of the roots of its denominator.

8. Prove that if x and a are positive and unequal, and m is any rational number greater than 1,

$$mx^{m-1}(x-a)>x^m-a^m>ma^{m-1}(x-a);$$

while if $0<m<1$ signs of the inequalities must be reversed.

[Suppose first that $a=1$ and let $m=p/q$. It follows from the inequality (3) of § 67 that, if ξ is any number greater than unity,

$$\xi^p-1 \gtrless (p/q)(\xi^q-1)$$

according as $p \gtrless q$, and it is easy to see, by similar reasoning, that the result remains true if $0<\xi<1$, though both sides of the inequality are then negative. Writing $x^{1/q}$ for ξ and m for p/q we obtain

$$x^m-1 \gtrless m(x-1) \qquad \text{(1)},$$

according as $m \gtrless 1$. If now we replace x by $1/x$, and multiply by $-x^m$, we obtain

$$mx^{m-1}(x-1) \gtrless x^m-1 \qquad \text{(2)}.$$

From (1) and (2) the result follows in the case of $a=1$. The proof may now be completed by writing x/a for x.]

9. Show that the inequality stated in Ex. 8 holds also if m is negative. Obtain corresponding inequalities when x and a are both negative. [See Chrystal's *Algebra*, vol. ii, pp. 43–45.]